

POSCO AI / BIGDATA ACADEMY 9TH 빅데이터 프로젝트

# 행복그린 매장 고객 세분화 전략을 통한 매출 증대

[ A반 1조 ] 강민구 권수민 이용현 이정우 안태윤 정해유

# Contents

1. 추진 배경
2. 현상 및 개선 기회
3. 분석 계획
4. 분석 결과
  - 4.1 RFM 분석
  - 4.2 제품 연관규칙 분석
  - 4.3 신규 활동고객 분석
  - 4.4 시간에 따른 매장별 매출 분석
5. 개선안 적용 방안
6. Learned Lesson
7. 첨부

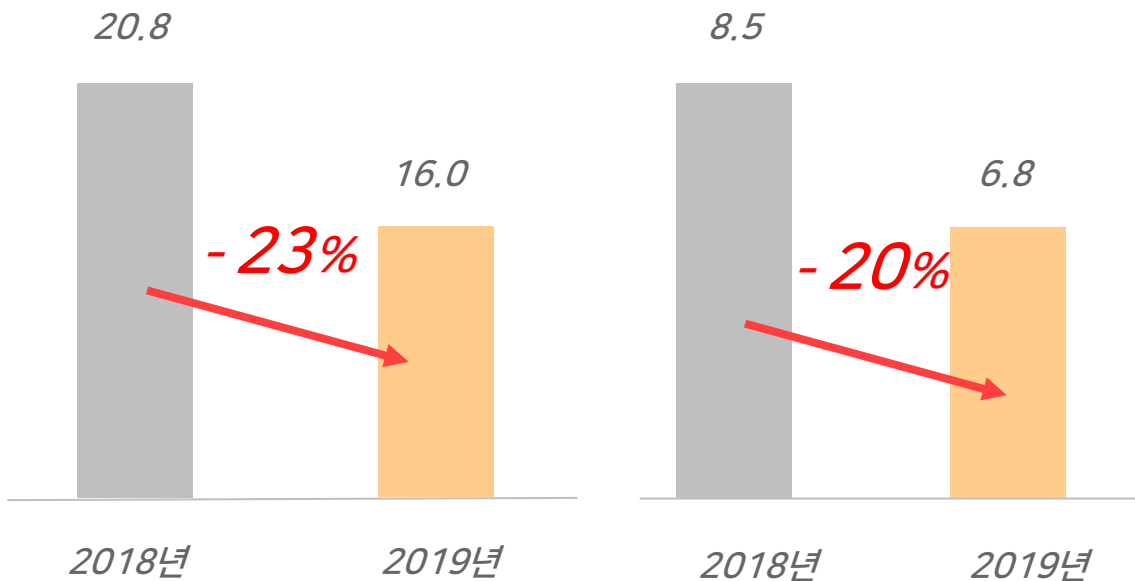
## 추가 분석

1. RFM을 이용한 추가 분석
  - 1.1 RFM에 따른 매장 특성 파악
  - 1.2 Rodgers 모형
  - 1.3 K-means clustering
  - 1.4 결과 비교 및 프로모션 전략
2. 신규고객의 RFM 특성 파악
3. 유사한 상품을 찾기 위한 새로운 방법 시도
  - 2.1 item2vec을 이용한 유사 상품 찾기
  - 2.2 프로모션 전략 수립

## 1. 추진배경

### 19년도 매출 급락에 대한 대책과 강화된 회원 관리를 통한 경영 혁신이 필요

[ 내부 상황 ] 18년 대비 19년도 매출이 20% 하락



연도별 판매 수량 (만 개)

연도별 매출 (억 원)

[ 외부 상황 ] 회원제를 통한 충성 고객 확보 경쟁이 치열

미끼 상품을 통해 '체리피커' 고객의 방문만을 유도했던 유행을 벗어나,

오직 회원에게만 혜택을 주는 회원제를 통해 지속적이고 안정적인 '충성 고객'을 확보하는 전략을 사용하고 있음.

파격적인 혜택을 통한 충성 고객 확보 경쟁은 앞으로도 계속될 것으로 예상됨.

※ 체리피커란? : 이용 실적은 좋지 않으면서 실속만 챙겨가는 고객

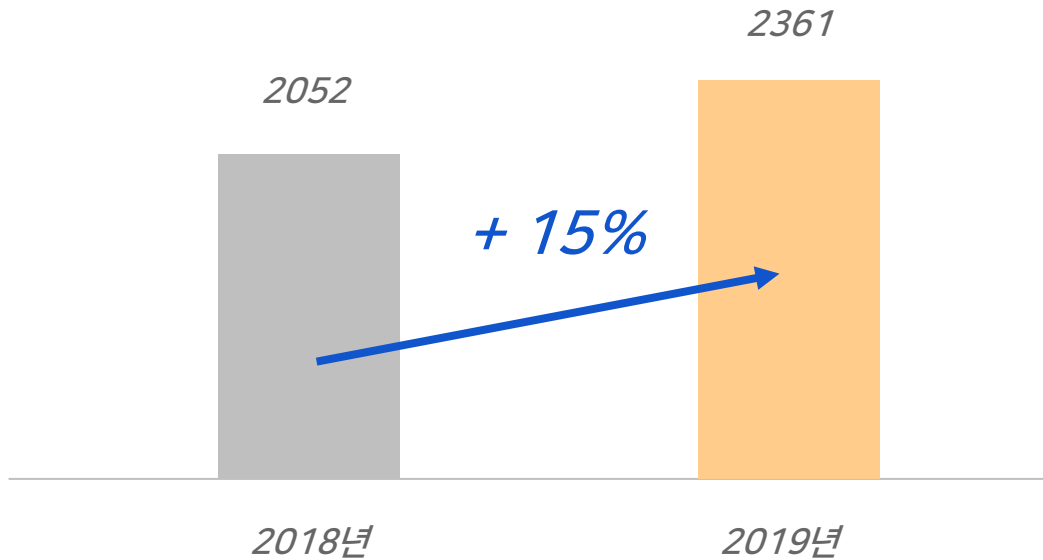
<회원제 강화하는 유통업계, '체리피커'보다 '충성고객'>

## 2. 현상 및 개선기회

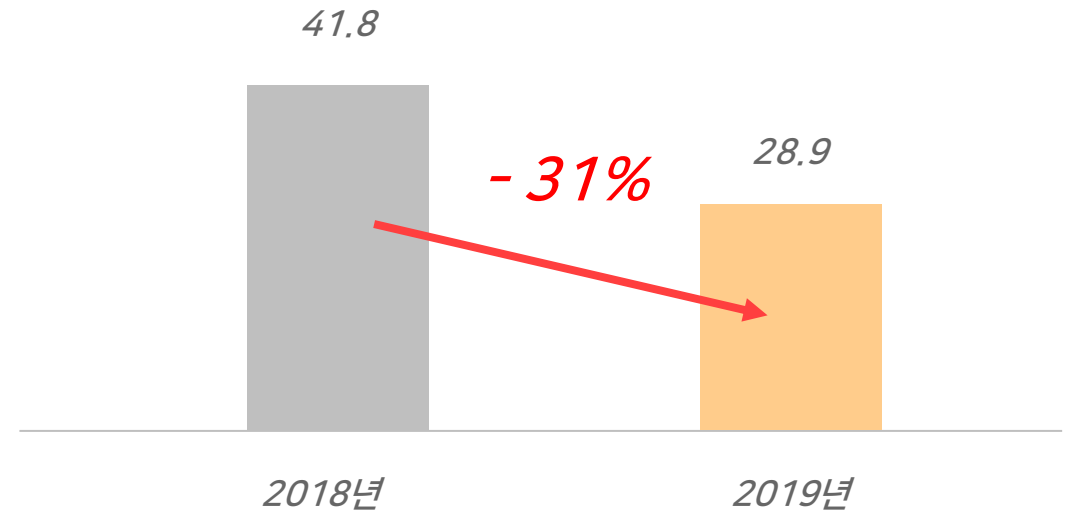
전년대비 활동 고객 수는 15%<sup>(309명)</sup> 증가. 그러나 1인당 소비 금액은 31%<sup>(13만원)</sup> 감소

연간 활동 고객 수 (단위 : 명)

※ 활동 고객이란? : 해당 년도에 1번 이상 구매를 한 고객



1인당 총 소비 금액 (단위 : 만 원)



## 2. 현상 및 개선기회

### [ 목 표 ] 2020년도 매출 8.5억원 (25%) 달성

매출에 영향을 미치는 핵심적 요소(Key Performance Indicator)를 다음으로 설정한다.

핵심 성과 지표	현수준 (19년)	목표 수준
객단가	289,399원	294,000원
이용 고객 수	2,361 명	2,900명

※ 객단가란? 총 매출 / 이용 고객 수

## 2. 현상 및 개선기회

### 고객 관리 / 구매 유도 / 매장 운영 측면에서 개선 기회 도출

항목	데이터 확보	데이터 활용 및 개선 예시	성과지표
고객 관리	고객 등급 및 매출액	활동 고객의 등급 재조정 및 고객 관리 이탈 고객 관리	이용 고객수 객단가
구매 유도	고객별 구매 데이터	신규 고객의 최초 구매 유도 관련된 제품 분석(연관분석)을 통해 관심이 적은 제품 구매 유도	
매장 운영	시간별 구매 데이터	매장 운영 프로세스 개선	

### 3. 분석계획

- 매출 하락 원인 분석 후, 다양한 분석 모델을 통해 매출을 상승시키기 위한 프로모션 전략 수립

	분석 대상	분석 방법	개선안 도출
고객 관리	고객 등급 구조	탐색적 분석 RFM 분석	등급별 고객 관리 제품 추천 방안 제시 운영 프로세스 개선
구매 유도	신규고객 고객 / 제품	의사결정나무 랜덤 포레스트 연관규칙 분석	
매장 운영	시간 / 매장 / 제품	탐색적 분석	



### 3. 분석계획 - 데이터 전처리

- '구매 수'와 '총 지불한 금액'을 제외하고 특별한 이상치나 결측치는 발견되지 않음.
- 따라서 총 144개의 결측치를 제거함

구매 수는 0이나 지불한 금액이 있는 26개 데이터에 대해서 제거를 진행함

(단위: 원)

대분류	중분류	소분류	제품명	구매 수	단가	총 지불한 금액
간식	과자	한과/유과	유과꾸러미1(1.2kg)	0	47,000	9,400
간식	과자	한과/유과	유과꾸러미1(1.2kg)	0	47,000	9,400
간식	과자	한과/유과	유과꾸러미1(1.2kg)	0	47,000	15,000
채소	잎채소	아스파라거스	아스파라거스(150g)	0	4,150	2,905

[ 구매 수가 0개 인 구매 데이터의 일부 ]

### 3. 분석계획 - 데이터 전처리

- '구매 수'와 '총 지불한 금액'을 제외하고 특별한 이상치나 결측치는 발견되지 않음.
- 따라서 총 144개의 결측치를 제거함

총 지불 금액이 0원인 데이터 118개에 대해서 제거를 진행함

(단위: 원)

대분류	중분류	소분류	제품명	구매 수	단가	총 지불한 금액
채소	잎 채소	포기생채	포기생채(2포기)	1	1,400	0
채소	뿌리채소	무	무(1kg이상)	1	1,800	0
축산물	육가공	돼지가공	꼬마소시지(300g)	1	4,200	0
간식	떡	떡	카스텔라 인절미(180g)	1	3,000	0

[ 총 지불한 금액이 0원 인 구매 데이터의 일부 ]

## 분석결과

4.1 RFM 분석

4.2 제품 연관규칙 분석

4.3 신규 활동고객 분석

4.4 시간에 따른 매장별 매출 분석

## 4.1 RFM 분석의 목적

1. 잘못 분류된 고객 등급 재조정
2. 2018년에서 2019년으로의 등급 이동에 따른 고객 세분화

## 4.1 RFM 분석 (1) 기존 등급 재조정 필요

등급별 1인 매출액을 살펴보았을 때, 기존 등급이 잘못 분류되어 있다는 사실을 알 수 있음. 그러므로 고객 등급 재조정이 필요함

- Platinum 의 1인 최대 매출액(4,488,260원)이 Royal의 객단가(1,947,640원) 보다 높음
- Platinum 의 1인 최소 매출액(1,300원)이 Gold 의 1인 최소 매출액(1,300원)과 같음
- 이를 통해 기존 등급제도가 고객을 잘못 분류하고 있다는 사실을 알 수 있음
- 잘못된 고객 분류는 고객 불만족으로 이어질 수 있음

(단위: 원)

등급	ROYAL	PLATINUM	GOLD	SILVER
1인 최대 매출액	25,150,240	4,488,260	3,061,650	1,520,670
1인 최소 매출액	8,000	1,300	1,300	1,000
1인 매출액 (중앙값)	1,586,435	275,200	59,425	18,100
객단가	1,947,640	409,520	158,621	39,780

[ 2019년 등급별 1인 매출액 비교 ]

(\*19년도 거래고객 기준)

## 4.1 RFM 분석 (1) 기존 등급 재조정 필요

기존 등급을 어떻게 분류하였는지에 대한 자료는 없으나  
회원 정보를 살펴보면, 약간의 부적절함을 발견할 수 있음

고객 ID	기존 등급	2018-12-31로부터 n일 전에 방문	2018년 총 방문 횟수	2018년 총 소비 금액
C418404	Gold	2	70	1,581,000
C404906	Gold	4	79	1,821,460
C422720	Gold	5	113	4,075,630
C541996	Gold	3	84	1,866,885
C472947	Gold	2	70	1,914,550

3일에 한번 꼴로 방문하여  
1년간 400만원을 썼음에도 Gold 등급

고객 ID	기존 등급	2018-12-31로부터 n일 전에 방문	2018년 총 방문 횟수	2018년 총 소비 금액
C222499	Royal	75	1	20,350
C354460	Royal	145	1	21,200
C450409	Royal	107	4	67,100
C450436	Royal	187	1	15,450
C364526	Royal	80	4	69,100

딱 1번 방문하여 1.5만원을 썼음에도  
Royal 등급

## 4.1 RFM 분석 (2) RFM 분석을 통한 새로운 등급 체계 구축

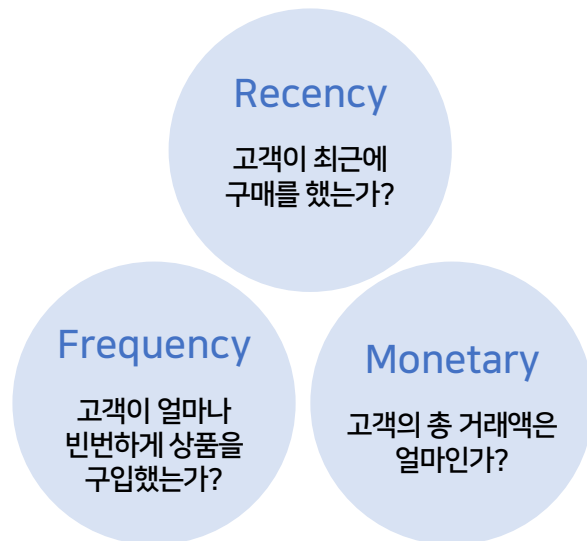
- RFM의 정의 및 RFM 등급 산정 방법

(RFM 분석의 정의)

- 최근 구매월, 구매 빈도, 구매 금액을 고려하여 고객의 가치를 평가하는 방법
- R/ F/ M 을 총 6등급으로 분류하고, 가중치를 부여하여 총 RFM 점수를 산정

(RFM 등급 산정 방법)

- Rogers의 모형을 활용하여 통계분포를 사용. R, F, M 의 분포를 고려하여 6개 등급으로 분류함
- R은 연도별 고객 등급 변화를 분석하기 위해 12개월로 설정



등급	누적구간
1	0 ~ 2.5%
2	2.5 ~ 16%
3	16 ~ 50%
4	50~84%
5	84 ~ 97.5 %
6	97.5 ~ 100%

[ \*참고자료: RFM에서 등급 부여 방법에 관한 연구(류귀열, 문영수/2013) ]

## 4.1 RFM 분석 (2) RFM 점수 산출 방법 및 등급 산출 예시

- R, F, M 등급의 1% 가 매출에 기여하는 정도를 가지고 각각의 가중치(w)를 구함
- RFM 점수는 아래의 공식을 따름
  - R : 19/12/31 을 기준으로 가장 최근에 방문한 날을 뺀 날 수  
EX) C541748 고객의 경우 19/12/31에 가장 마지막으로 방문했으므로 0일이다.
  - F : 19/01/01 - 19/12/31 의 기간 동안 해당 고객이 방문한 횟수
  - M : 19/01/01 - 19/12/31 의 기간 동안 해당 고객이 거래한 총 거래액 (고객 기준 객단가)

$$\text{RFM 점수} = w1 \times \text{R점수} + w2 \times \text{F점수} + w3 \times \text{M점수}$$

R의 가중치(w1)  $\div$  0.055 , F의 가중치  $\div$  0.74 , M의 가중치  $\div$  0.205

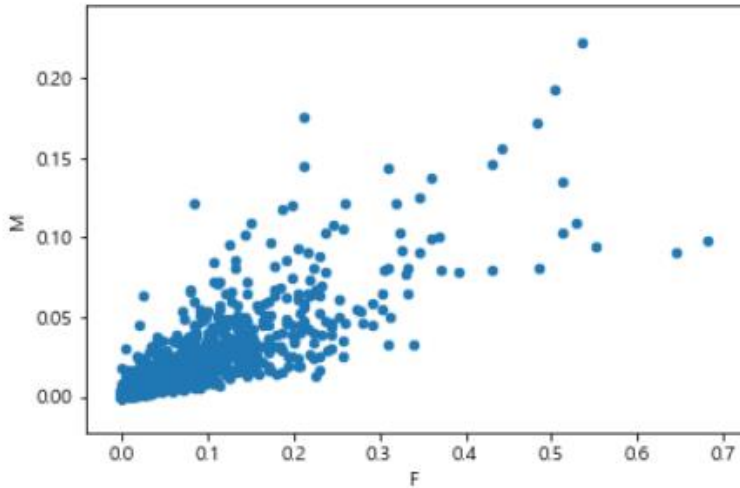
Customer ID	R	F	M	R 점수	F 점수	M 점수	RFM 점수	등급
C541748	0	289	25,144,440	6	6	6	6	ROYAL
C238894	27	21	536,100	5	3	4	3.32	GOLD



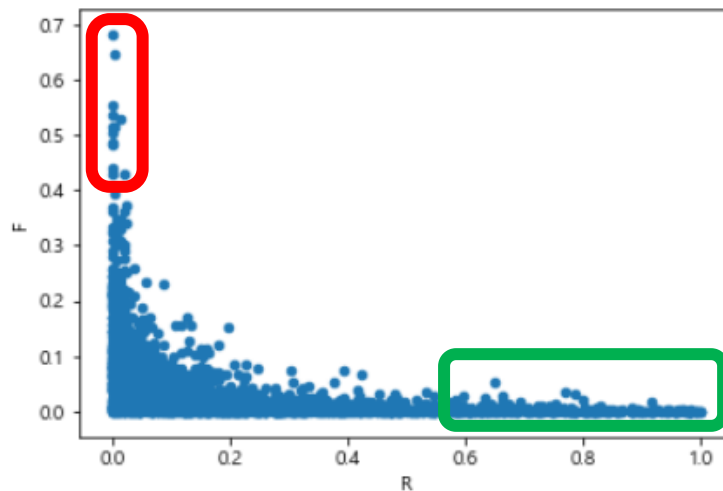
## 4.1 RFM 분석 (1) R,F,M 분포 확인

자주 방문하여 많은 금액을 쓴 **충성 고객들**은 최근에도 방문하였음. 즉 이탈 고객이 아님.  
**이탈 위험 고객들**은 대부분 자주 방문하지 않았으며 소비 금액도 적음.

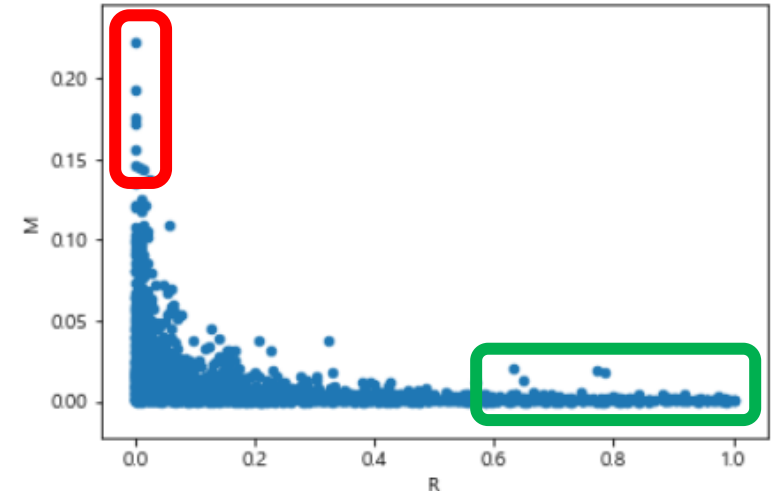
F - M 에 대한 그래프



R - F 에 대한 그래프



R - M 에 대한 그래프



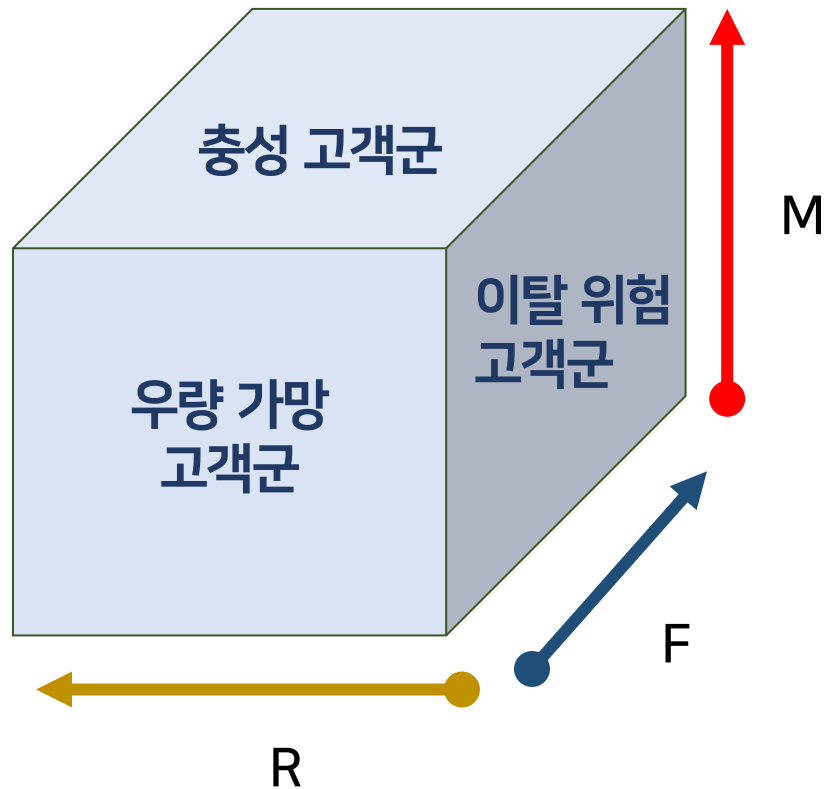
- **충성 고객**이 이탈한 경우는 없는 것으로 파악됨.
- **최근성 점수 R이 낮은 집단**의 경우, 2018년 12월 31일로부터 마지막 방문 날짜까지의 간격이 크기 때문에 이미 이탈했거나, 이탈할 위험이 큰 고객으로 볼 수 있음.

\*MinMax scaling한 결과를 plot하였음  
\*R의 경우 작을 수록 최근에 방문

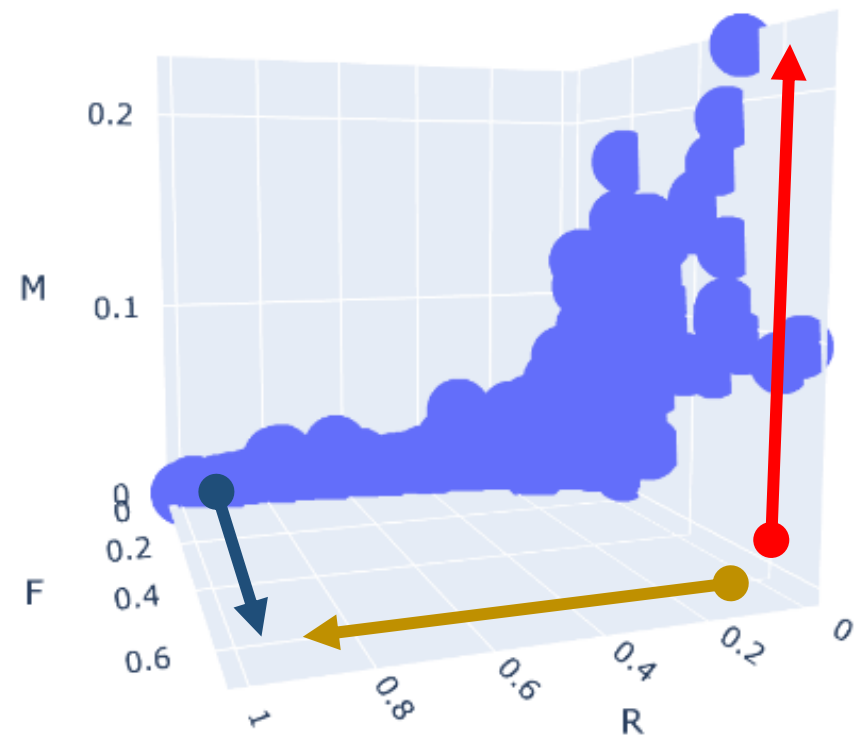
## 4.1 RFM 분석 (1) R,F,M 분포 확인

**충성 고객군이 부족하며 이탈 위험 고객군이 다수 존재함.**  
**우량 가망 고객군을 충성 고객으로 만들 전략이 필요함**

R - F - M 별 고객 특성



R - F - M 에 대한 그래프



\*MinMax scaling한 결과를 plot하였음  
\*R의 경우 작을 수록 최근에 방문

## 4.1 RFM 분석 (2) R,F,M 점수를 통해 등급을 나누는 여러가지 방법

최종 보고서에서는 '3.Rodgers의 모형'을 이용하였으나  
본 분석에서는 **고객 특성에 따른 세밀한 분류를 위해 K-means clustering을 이용**하고 **두 결과를 비교**한다.

### 1. 20%씩 잘라 등급을 부여한다.

- 그러나 획일적으로 나누기 때문에 R,F,M 각각의 분포가 무시된다. 점수 차이가 큼에도 같은 등급으로 분류될 가능성이 있다.

### 2. K-means 등의 군집 분석을 실시한다.

- 분포를 고려할 수 있다. 그러나 군집별 특징을 해석하기 어려울 수 있다.

=> R,F,M 평균을 이용하여 군집별 특성을 해석해본다면 유의미한 결과를 도출할 수 있을 것으로 기대된다.

### 3. Rodgers 모형을 사용한다.

- R,F,M 각각의 분위수를 이용하여 등급을 나눈다. 분포를 고려하면서도 해석이 쉽다.

=> 해석은 쉽지만 R,F,M을 선형 결합한 결과로 등급을 분류하기 때문에 고객의 특성을 모두 고려한 등급이라고 말할 수 없다.

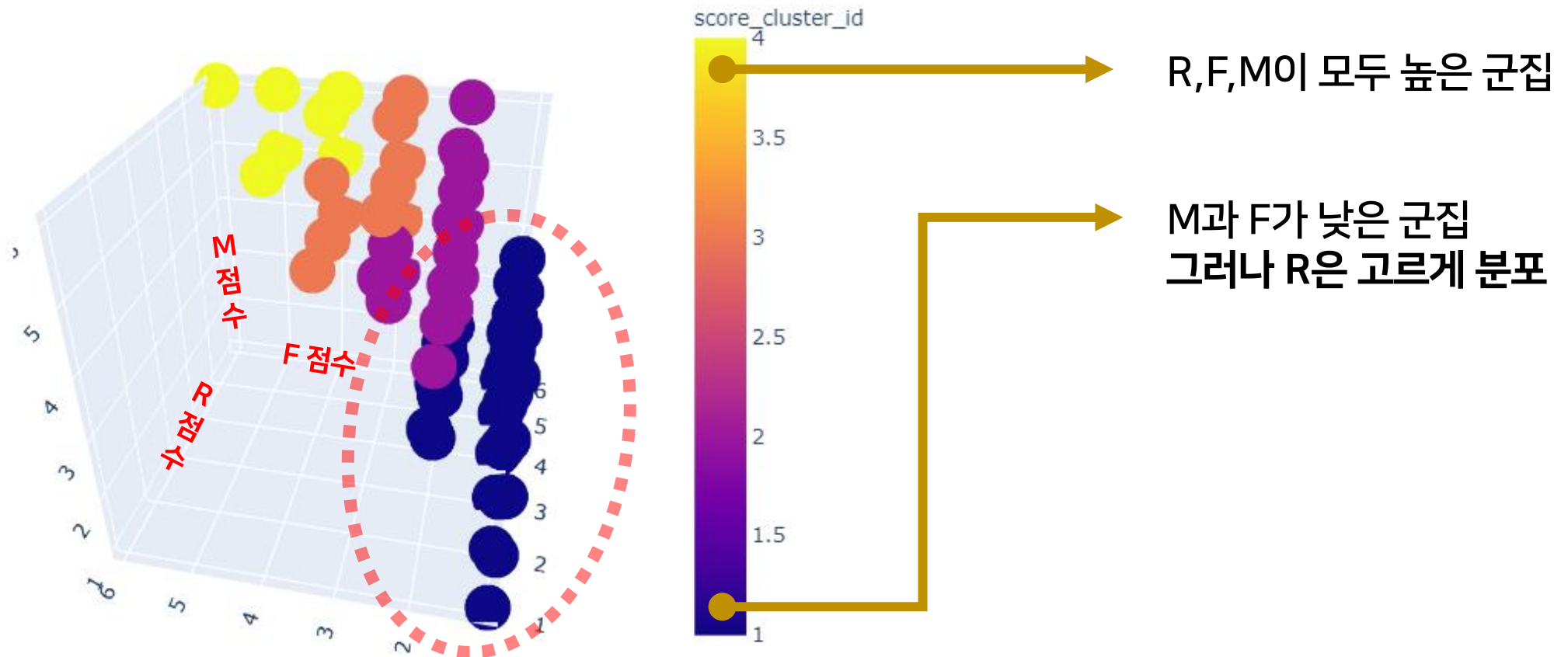
Ex) F는 높고 M이 낮은 고객과, F는 낮고 M이 높은 고객이 같은 RFM 점수를 가진다면, 고객 특성이 다름에도 같은 등급에 속할 수 있다.

## 4.1 RFM 분석 (3) Rodgers 모형에 의한 등급 분류

### 1. Rodgers 모형에 의한 등급 분류

F에 대한 가중치가 가장 크기 때문에 F 점수에 의해 등급이 분류된 형태를 띤다.

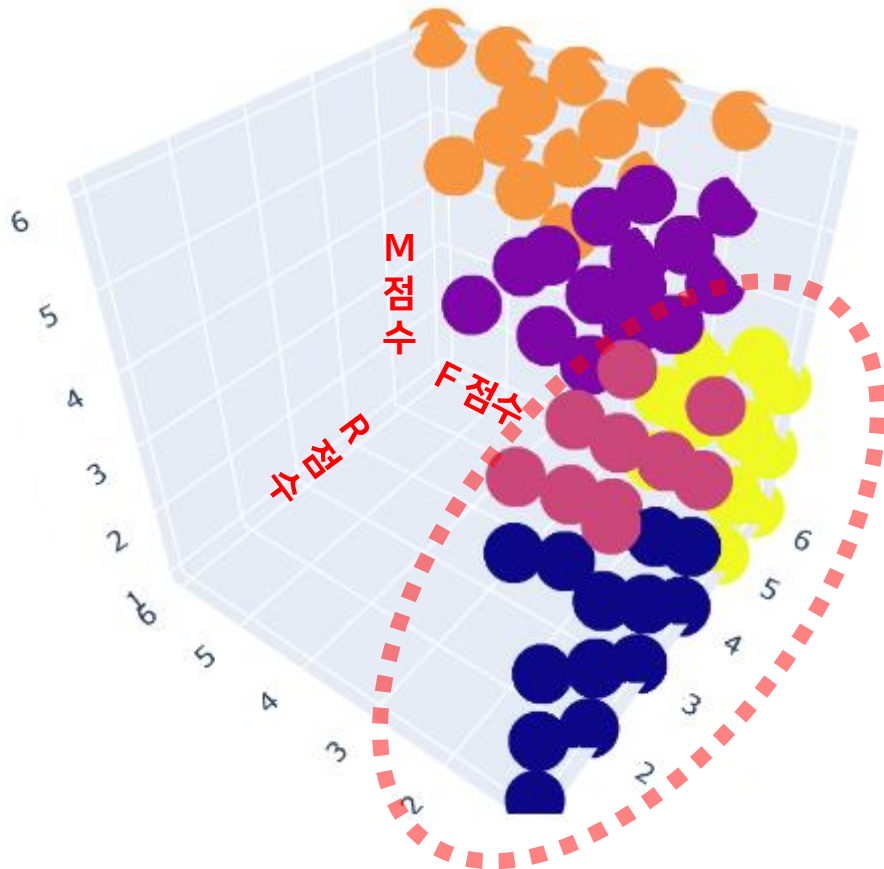
가장 낮은 등급의 경우, R 특성을 고려하지 못하고 있다.



## 4.1 RFM 분석 (3) Rodgers 모형에 의한 등급 분류

### 2. K-means에 의한 등급 분류

F가 낮은 등급이 **R 특성을 반영하여 더 세밀하게 분류**되었음을 확인할 수 있다.  
=> **고객 특성별 프로모션**을 위해서는 K-means 에 의한 군집화가 더 적절하다.



score\_cluster\_id



F, M은 낮지만 R은 높은 군집

R, F, M이 모두 낮은 군집

## 4.1 RFM 분석 (3) Rodgers 모형에 의한 등급 분류

### 2. K-means에 의한 등급 분류

군집별 특성을 중앙값을 통해 확인하고 고객 특성을 분류하면 다음과 같다.

군집	R 점수	F 점수	M 점수	군집 특성
자주 방문하지도 않으며 최근에 방문하지도 않은 고객	3	1	2	이탈 위험 고객
최근에 방문했지만 자주 오지 않아 소비 금액이 적은 고객	4	1	2	관심 고객
일반 고객	4	2	3	일반 고객
최근에 방문했으며 많이 소비했지만 자주 오지는 않는 고객	5	2	4	충성 고객으로 전환 가능한 고객
R,F,M이 모두 높은 충성 고객	6	3	5	충성 고객

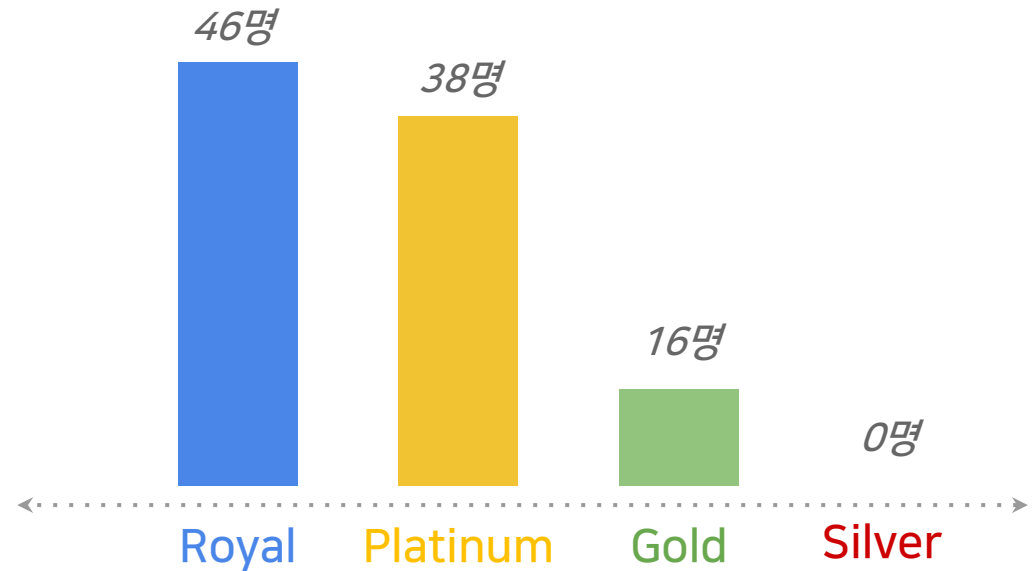
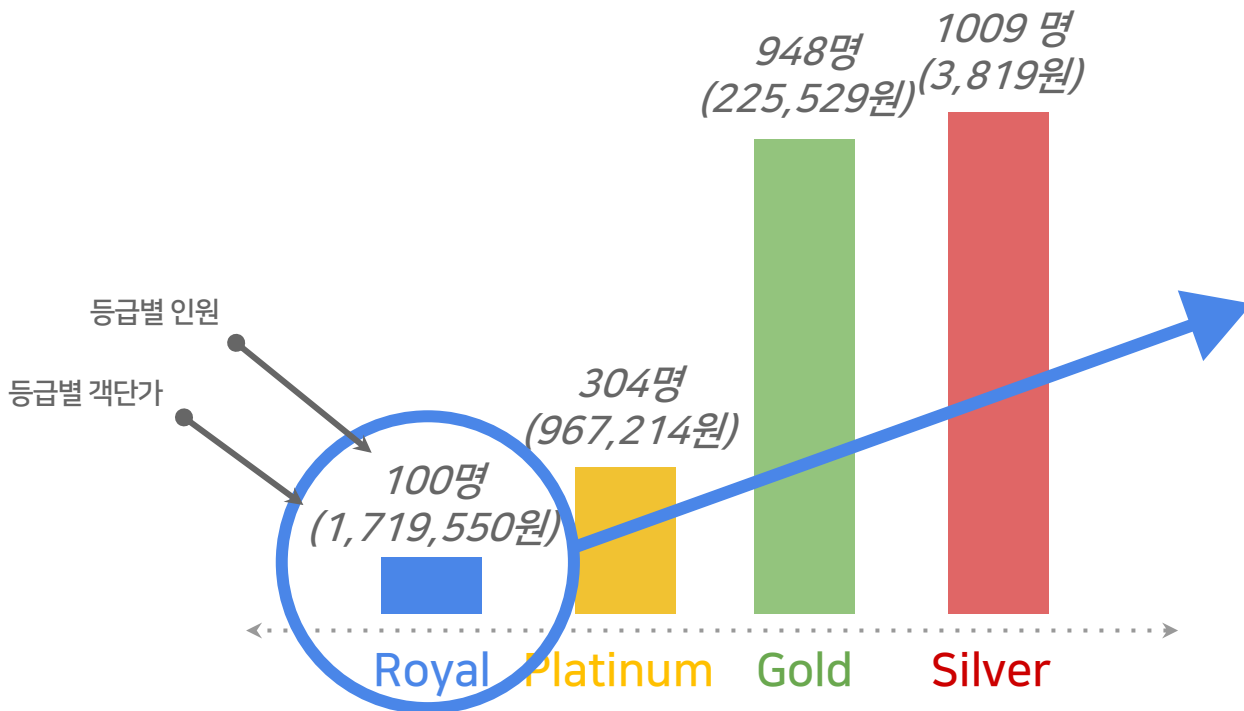
충성 고객으로 전환 가능한 고객군 역시 F 점수가 부족(즉 자주 방문하지 않음)하므로 군집 특성을 이용하여 전반적인 고객 유지 전략 + 적극적인 이탈 방지 전략을 세울 필요가 있음.

## 4.1 RFM 분석 (3) 새로운 등급 구조

- RFM 분석을 통해 기존에 잘못 분류되어 있던 고객을 알맞은 등급으로 재조정함
- 따라서 2018년 대비 2019년 고객 등급 이동 추이를 파악할 수 있게 되었고 고객 세분화 전략을 구상할 수 있게 됨
- \*오른쪽 그래프를 보면 **Platinum** 이나 **Gold** 에 있던 고객들이 등급 상향화가 되었음을 알 수 있음

새로운 등급의 등급별 인원 및 객단가

\*새로운 등급에서 **Royal**에 속한 고객의 기존 등급



## 4.1 RFM 분석 (4) 2018에서 2019년으로의 고객 등급 이동

18'년의 상위 등급 고객층(Royal, Platinum)은 상당수(71%)가 19'년에도 상위 등급을 유지  
반면 18'년의 하위 등급 고객(Gold, Silver)이 19'년에 상위등급으로 이동한 경우는 적음(6.9%)  
**Silver** 등급의 경우 고객 이탈률이 37.7% 로 이탈률이 가장 높음

18년 등급	19년 등급	고객수
Royal	Royal	64
	Platinum	21
	Gold	4
	Silver	1
Platinum	Royal	21
	Platinum	143
	Gold	91
	Silver	3
	★이탈	1

18년 등급	19년 등급	고객수
Gold	Royal	8
	Platinum	101
	Gold	454
	Silver	157
Silver	★이탈	36
	Platinum	8
	Gold	204
	Silver	378
	★이탈	357

등급 상승

등급 유지

등급 하락

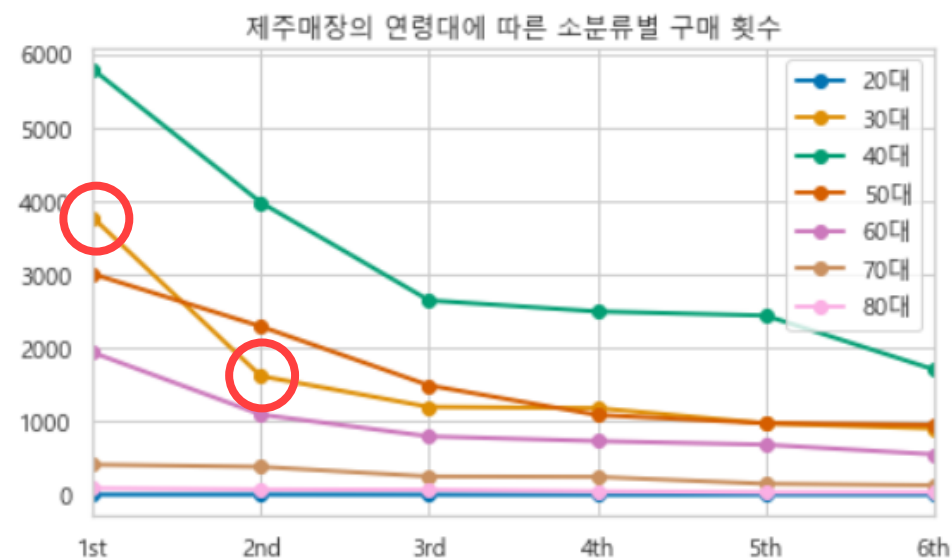


## 4.2 연관규칙 분석의 목적

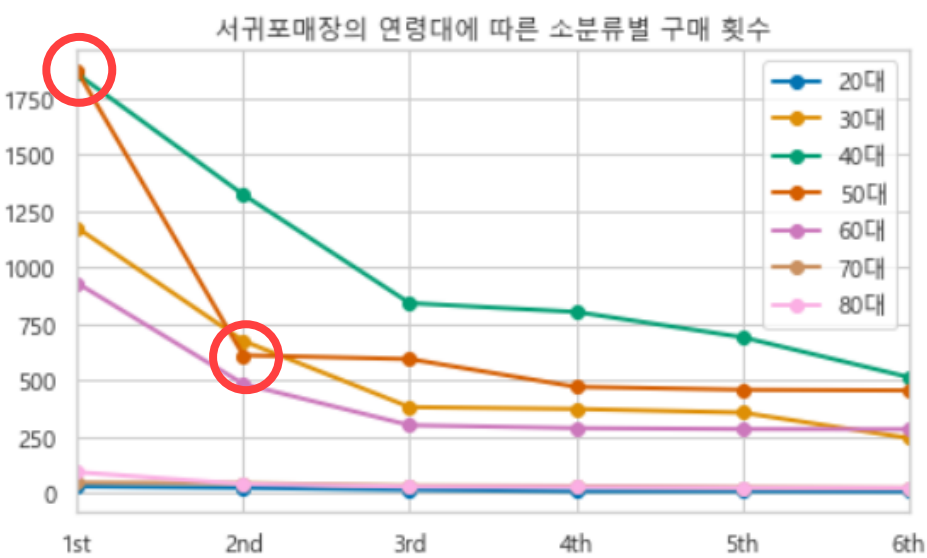
선호도가 높은 특정 제품과 관련이 있는 제품을 탐색

## 4.2 제품 연관규칙 분석 (1) 연관규칙 분석을 위한 대상 설정

특정 제품군에 대한 선호가 두드러진 연령대를 선정하여 다른 제품군과의 연관분석을 실시  
선호도가 높은 제품을 이용하여 다른 제품의 구매를 유도



	20대	30대	40대	50대	60대	70대	80대
1st	3	3774	5790	3005	1937	412	90
2nd	3	1616	3977	2291	1092	382	74

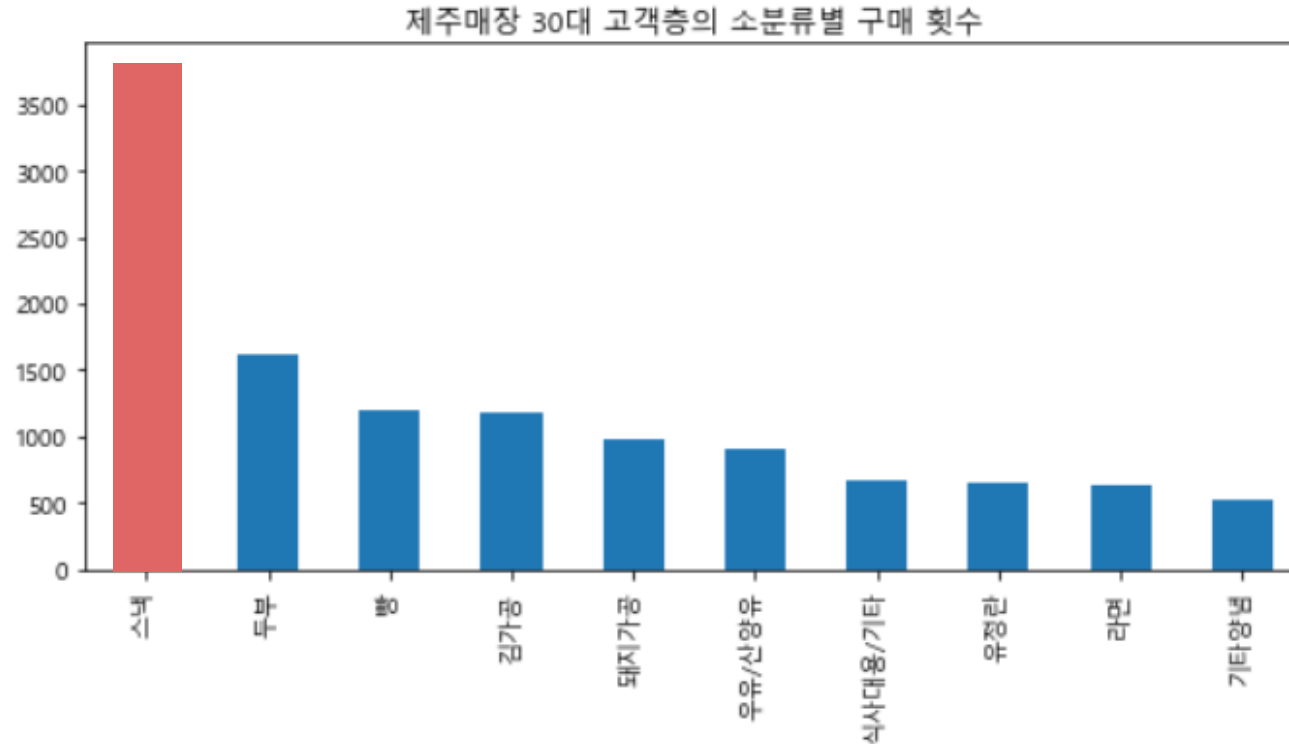


	20대	30대	40대	50대	60대	70대	80대
1st	30	1177	1895	1874	932	47	93
2nd	24	675	1323	611	484	44	40

※ 1st: 해당 매장, 해당 연령대에서 가장 많이 팔리는 제품  
2nd: 해당 매장, 해당 연령대에서 2번째로 많이 팔리는 제품  
6th: 해당 매장, 해당 연령대에서 6번째로 많이 팔리는 제품

## 4.2 제품 연관규칙 분석 (1) 연관규칙 분석을 위한 대상 설정

제주 매장 30대 고객층에서는 '스낵' 제품을 이용하여 다른 제품을 함께 구매하도록 유도

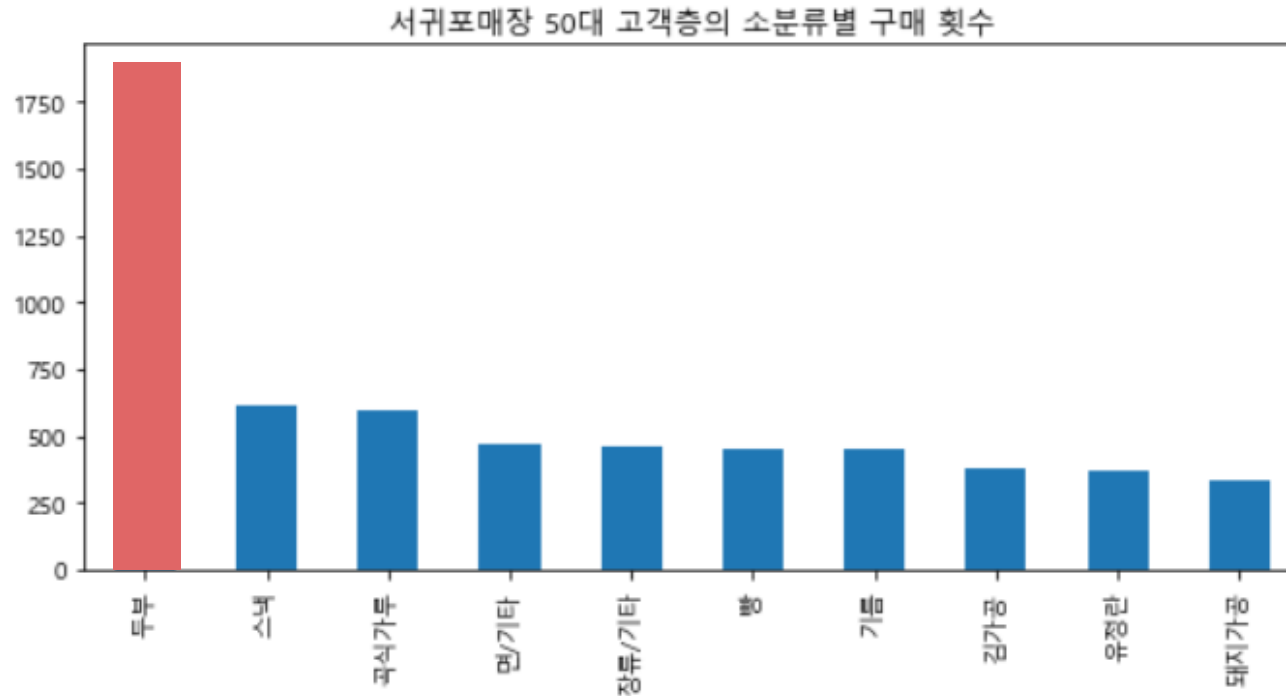


소분류	구매 횟수
스낵	3774
두부	1616
빵	1196
김가공	1180
돼지가공	973
우유/산양유	903

- '스낵' 제품이 다른 제품들에 비해 상대적으로 큰 비중을 차지
- 연관규칙 분석을 활용하여 제주매장 30대 고객들이 '스낵'을 구매할 때 어떤 제품을 함께 구매하는지 확인해볼 필요가 있음

## 4.2 제품 연관규칙 분석 (1) 연관규칙 분석을 위한 대상 설정

서귀포매장 50대 고객층에서는 '두부' 제품을 이용하여 다른 제품을 함께 구매하도록 유도



소분류	구매 횟수
두부	1874
스낵	611
곡식가루	594
면/기타	471
장류/기타	458
빵	455

- '두부' 제품이 다른 제품들에 비해 상대적으로 큰 비중을 차지
- 연관규칙 분석을 활용하여 서귀포매장 50대 고객들이 '두부'를 구매할 때 어떤 제품을 함께 구매하는지 확인해 볼 필요가 있음

## 4.2 제품 연관규칙 분석 (2) 연관 규칙 선정 기준

유용한 연관 규칙을 선정한 기준은 다음과 같다.

항상도가 1보다 큰 규칙 중에서 신뢰도가 가장 높은 규칙을 선정한다.

\*신뢰도(confidence)란? 항목집합 X(antecedents)를 포함하는 거래 중에서, 항목집합 Y(consequents)도 포함하는 거래 비율

### 제주매장 30대

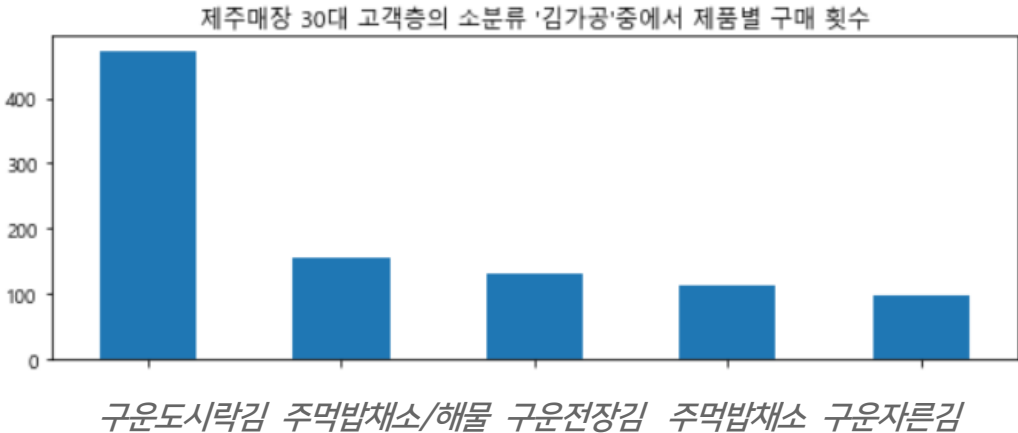
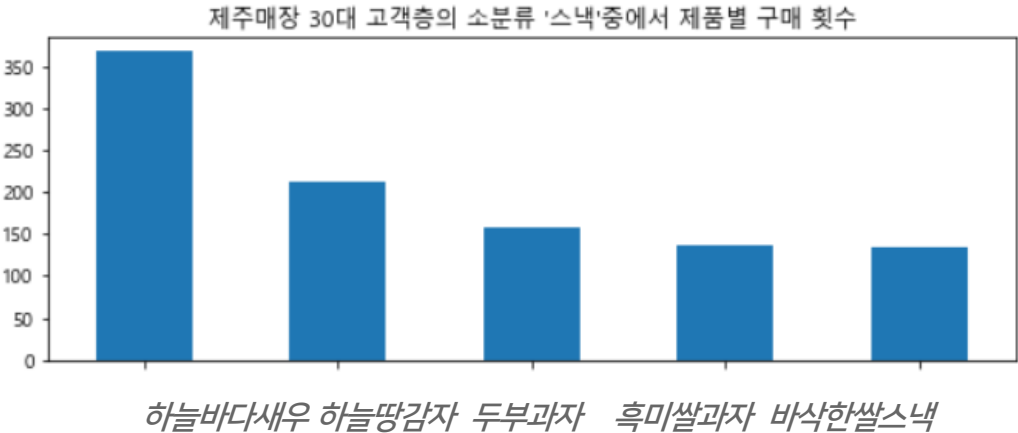
antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
(스낵)	(빵)	0.294330	0.129124	0.042912	0.145797	1.129125	0.004907	1.019519
(스낵)	(김가공)	0.294330	0.133119	0.050129	0.170315	1.279425	0.010948	1.044832
(스낵)	(돼지가공)	0.294330	0.110180	0.039820	0.135289	1.227886	0.007390	1.029037
(스낵)	(두부)	0.294330	0.190851	0.055670	0.189142	0.991047	-0.000503	0.997893
(스낵)	(우유/산양유)	0.294330	0.108247	0.035696	0.121278	1.120382	0.003835	1.014830

### 서귀포매장 50대

antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
(두부)	(곡식가루)	0.317570	0.120174	0.074187	0.233607	1.943910	0.036023	1.148009
(두부)	(곡식가루, 장류/기타)	0.317570	0.053145	0.040781	0.128415	2.416304	0.023904	1.086360
(두부)	(된장)	0.317570	0.063557	0.040781	0.128415	2.020459	0.020597	1.074414
(두부)	(면/기타)	0.317570	0.102169	0.071800	0.226093	2.212926	0.039355	1.160127
(두부)	(장류/기타)	0.317570	0.097397	0.056399	0.177596	1.823421	0.025469	1.097517
(두부)	(곡식가루, 면/기타)	0.317570	0.067896	0.054230	0.170765	2.515102	0.032668	1.124053

## 4.2 제품 연관규칙 분석 (2) 제주 매장 30대에 대한 연관규칙 분석

[ 제주 매장 30대 고객층 연관규칙 ] '스낵' → '소풍 / 나들이' → '김가공'

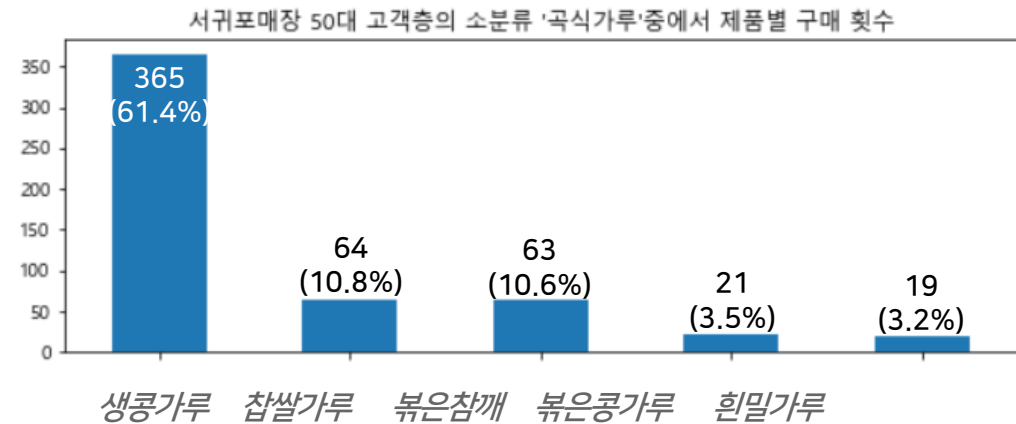
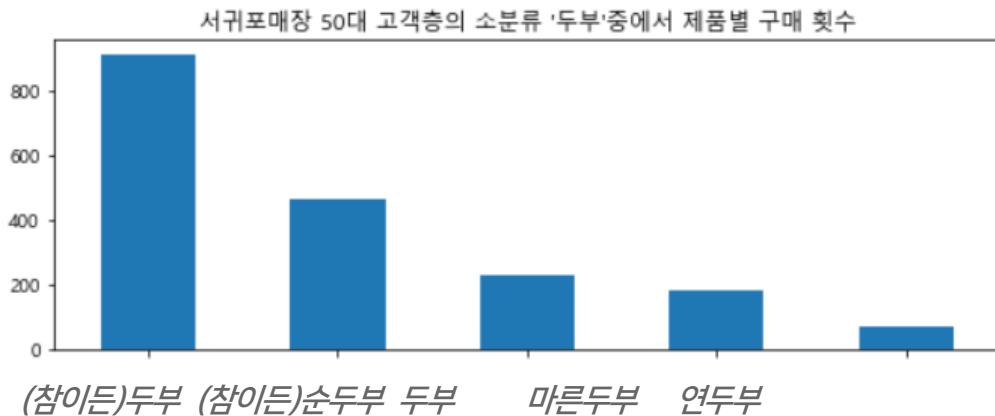


antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
(스낵)	(김가공)	0.294330	0.133119	0.050129	0.170315	1.279425	0.010948	1.044832

- '도시락김', '주먹밥채소', '전장김'과 같이 소풍, 나들이용 '김가공' 제품을 구매하고 있는 것으로 보아 소풍, 나들이를 위해 '스낵'을 구매하면서 '김가공' 제품을 함께 구매하는 것으로 생각해볼 수 있음

## 4.2 제품 연관규칙 분석 (2) 서귀포 매장 50대에 대한 연관규칙 분석

[ 서귀포 매장 50대 고객층 연관규칙 ] '두부' → '콩으로 만든 제품을 선호' → '생콩가루'



antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
(두부)	(곡식가루)	0.31757	0.120174	0.074187	0.233607	1.94391	0.036023	1.148009

- '생콩가루' 제품이 전체 곡식가루 구매 횟수 중 61.4%를 차지하고 있는 것으로 보아 콩으로 만든 제품을 선호하여 '두부'를 구매하면서 '곡식가루'의 '생콩가루' 제품도 함께 구매하는 것으로 생각해볼 수 있음

## 4.2.1 item2vec

item2Vec을 이용하여 함께 담긴 상품을 탐색



#### 4.2.1 item2vec (1) 순두부와 유사한 상품

**된장찌개나 청국장을 만들 때 순두부를 구매하는 것으로 추측할 수 있음**

순두부(참이든)과 함께 구매하는 상품

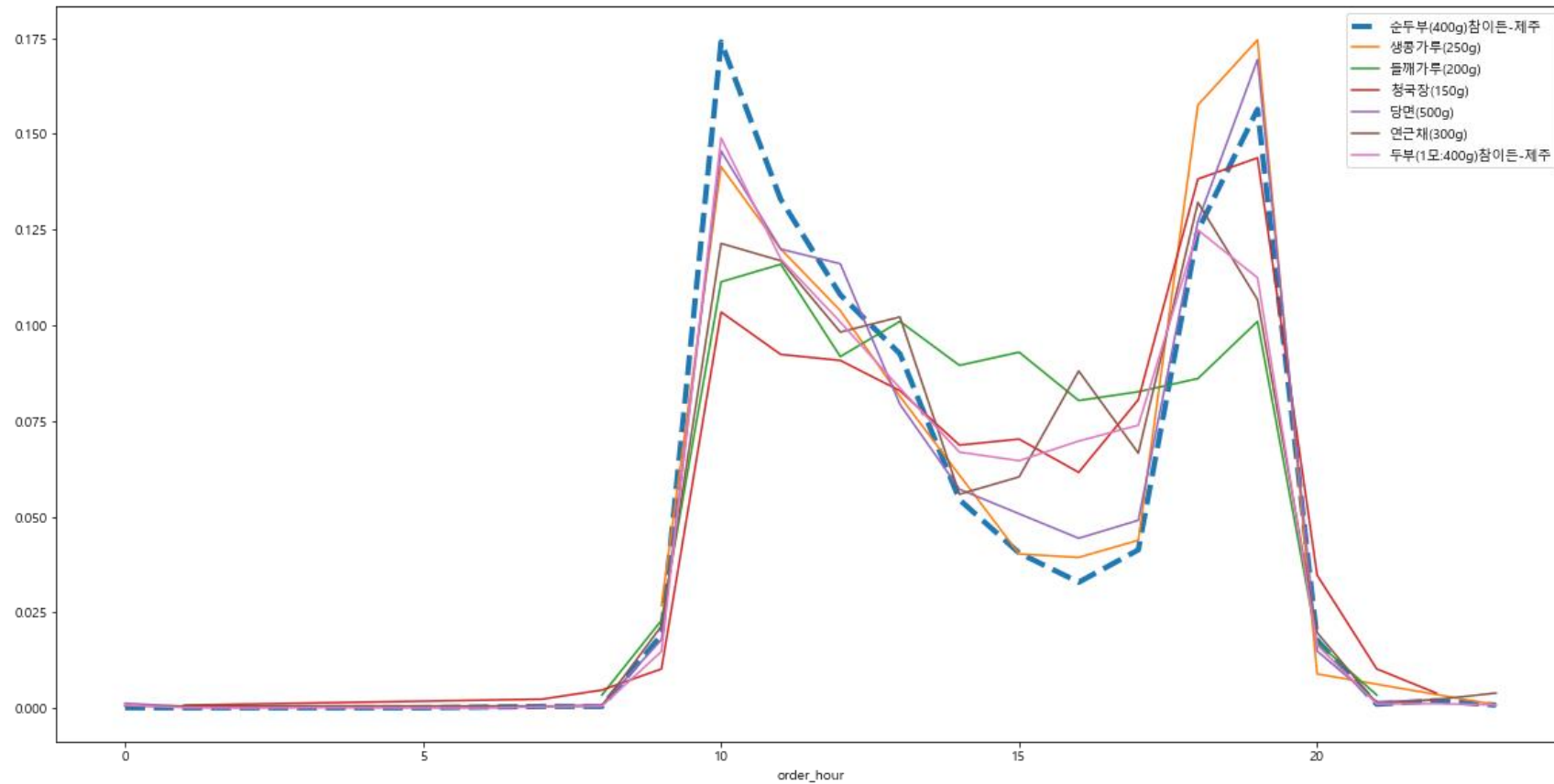
	제품	유사도
1	생콩가루(250g)	0.987
2	조선된장(900g)	0.978
3	들깨가루(200g)	0.978
4	들기름(330ml)	0.977
5	청국장(150g)	0.976
6	청국장(250g)	0.970
...		
9	두부(1모:400g)참이든	0.927

- 생콩가루/들깨가루  
- 국이나 죽을 끓일 때 고소해지도록 넣는 용도
- 조선 된장/청국장  
- 된장찌개, 된장국, 청국장 등의 재료가 됨

**=> [ 된장찌개/청국장 ]을 요리하지 않을까?**

## 4.2.1 item2vec (1) 순두부와 유사한 상품

순두부와 유사한 상품들의 판매 추이가 비슷하므로 구매가 증가하는 시간대에 프로모션 계획



- 순두부, 순두부와 유사한 상품들은 주로 오전 10시, 오후 7시에 수요가 급증함

#### 4.2.1 item2vec (2) 구운 전장김과 유사한 상품

나들이를 위해 도시락 재료를 구매하는 것으로 추측할 수 있음

구운 전장김과 함께 구매하는 상품

	제품	유사도
1	돈가스소스(240g)	0.981
2	오징어동그랑어묵(270g)	0.960
4	현미유(500ml)	0.955
5	단무지(반찬용:400g)	0.953
7	네모난햄(500g)-매장	0.952
8	우리밀빵가루(190g)	0.950
9	둥근긴어묵(270g)	0.945

- 전장김이란?  
자르지 않은 상태의 커다란 김.
- 1. 돈가스 소스. 우리밀 빵가루, 현미유  
- 돈가스를 만드는 재료로 추측됨
- 2. 긴 어묵, 네모난 햄, 단무지  
- 전장김과 함께 김밥 재료로 이용했을 것으로 추측됨
- 3. 원너 소시지, 동그랑 어묵  
- 소시지 볶음, 어묵 볶음 등 도시락에 자주 사용되는 반찬

=> [ 나들이 ] 를 계획하는건 아닐까?

#### 4.2.1 item2vec (2) 구운 전장김과 유사한 상품

**\* 빵가루와 등심/돈가스용 고기, 돈가스 소스의 유사도가 높음**

빵가루와 함께 구매하는 상품

	제품	유사도
8	등심/돈가스용500g	0.969
9	돈가스소스(240g)	0.969

- 빵가루의 용도는 다양하지만 등심/돈가스용 고기와 돈가스 소스가 함께 담겨 있다면 돈가스를 만들기 위한 용도로 추측할 수 있음
- 즉 구운 전장김과 돈가스 소스, 빵가루를 함께 구매한 것과 돈가스를 연결 지을 수 있음

## 4.2.1 item2vec (3) 요구르트와 유사한 상품

우유, 요구르트, 빵, 아이스크림 등은 함께 담길 가능성이 높으므로 함께 진열

블루베리 요거트와 함께 구매하는 상품

	item_	similarity
0	블루베리요구르트(120ml)	1.000
1	우유(범산/200ml)	0.917
2	★카스텔라(소:100g)/매장	0.917
3	소보로빵(75g)/매장	0.904
4	크루아상(150g)/매장	0.899
5	★크림빵(75g)/매장	0.897
6	★우유식빵(180g)/매장	0.891
7	구운유정란(6알)	0.890
8	소보루빵(75g)/매장	0.884
9	소보로단팥빵(90g)/매장	0.883
10	★유기농우유(종이팩/900ml)	0.881

아이스바우유와 함께 구매하는 상품

	item_	similarity
0	아이스바우유(75ml)	1.000
1	유기농블루베리아이스바(70ml)	0.976
2	아이스바딸기(75ml)	0.963
3	유기농요거트딸기아이스크림(100ml)	0.930
4	유기농플레인요거트아이스크림(100ml)	0.919
5	유기농아이스크림(기본형)	0.797
6	★크림빵(75g)/매장	0.788
7	크루아상(150g)/매장	0.764
8	소보루빵(75g)/매장	0.748
9	★콩콩크림빵(90g/매장)	0.741
10	양념닭갈비(500g)	0.738

크림빵과 함께 구매하는 상품

	item_	similarity
0	★크림빵(75g)/매장	1.000
1	★카스텔라(소:100g)/매장	0.983
2	크루아상(150g)/매장	0.976
3	소보로빵(75g)/매장	0.971
4	소보루빵(75g)/매장	0.953
5	★콩콩크림빵(90g/매장)	0.952
6	소보로단팥빵(90g)/매장	0.935
7	유기농플레인요거트아이스크림(100ml)	0.925
8	유기농아이스크림(기본형)	0.923
9	우유(범산/200ml)	0.922
10	★우유식빵(180g)/매장	0.919

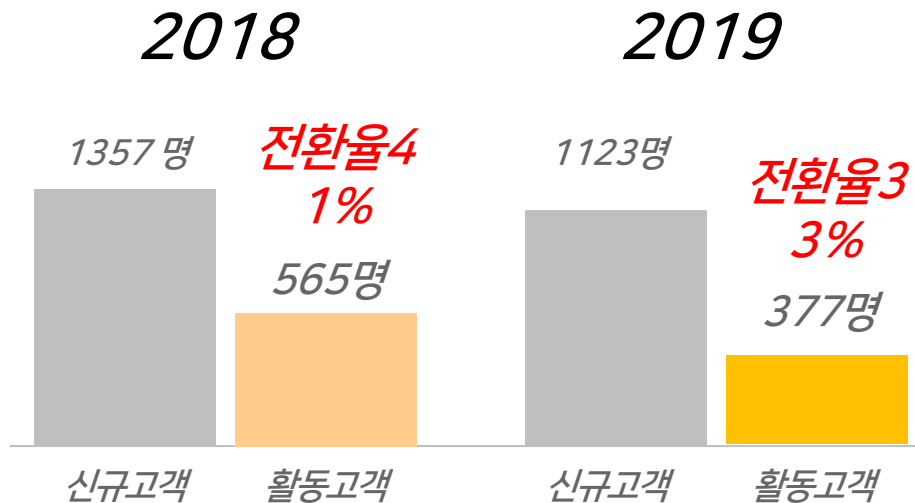
## 4.3 신규 활동고객 분석 목적

신규 활동고객들의 핵심 인자 탐색 및 선정

## 4.3 신규 활동고객 분석 (1) 신규 활동고객 전환율

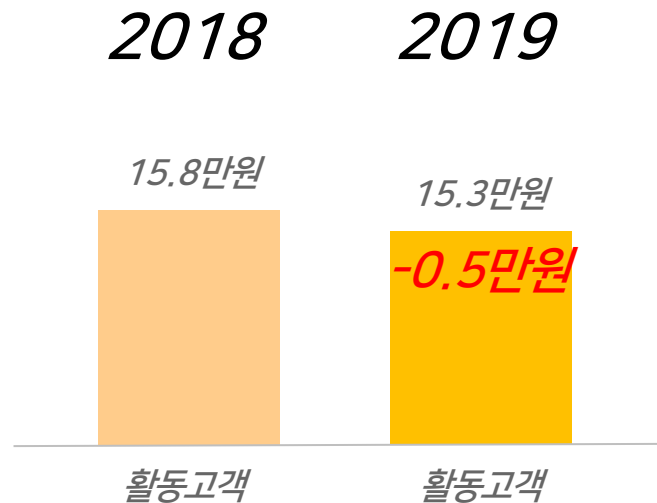
2019년 신규 고객의 활동고객 전환율은 33%에 불과,  
신규 활동고객 전환율을 높이는 것이 매출 증대의 핵심!

연도별 신규 고객 중 활동고객 수

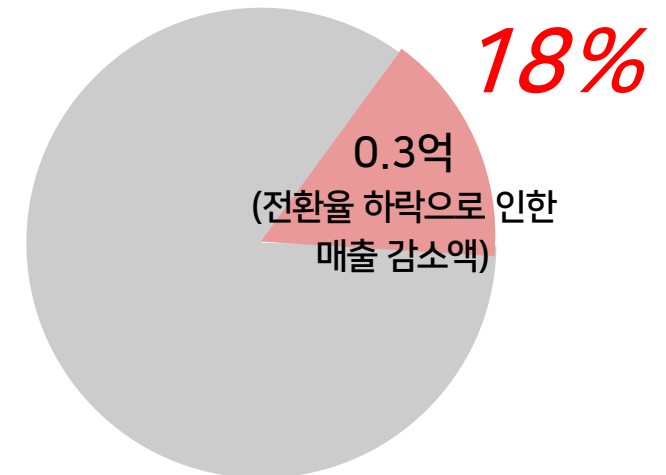


연도별 신규 활동고객 객단가

단위 : 만원



전환율 하락에 의한 매출감소 비율(19년)



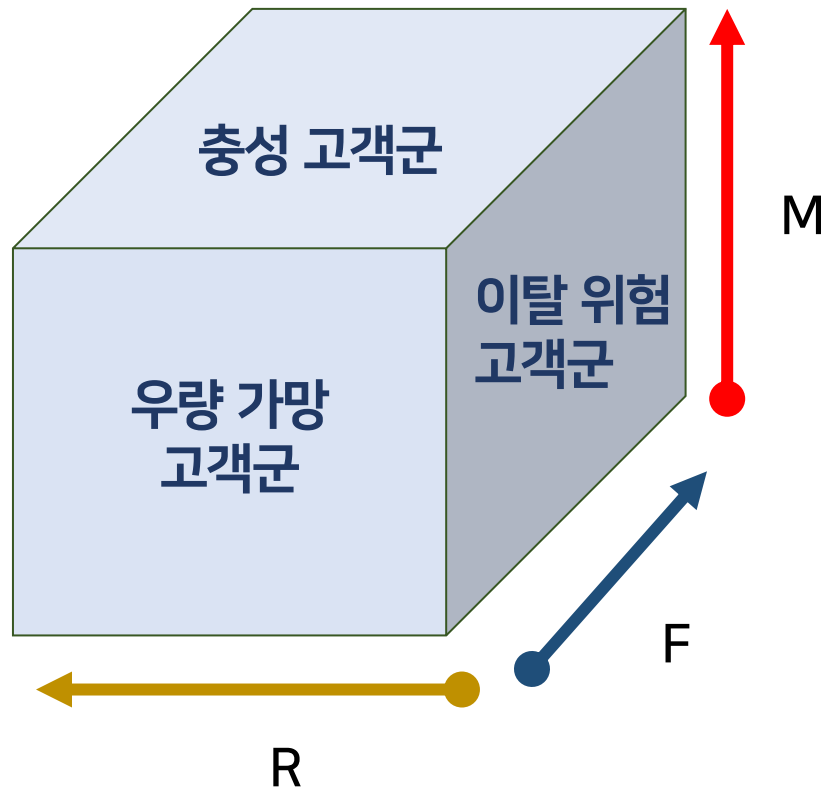
[ 전체 매출 감소액(1.7억) ]

- 연도별 신규 활동고객의 객단가 차이는 미비 (0.5만원)
- 신규 활동고객의 전환율을 높이면 매출증대를 예상할 수 있음

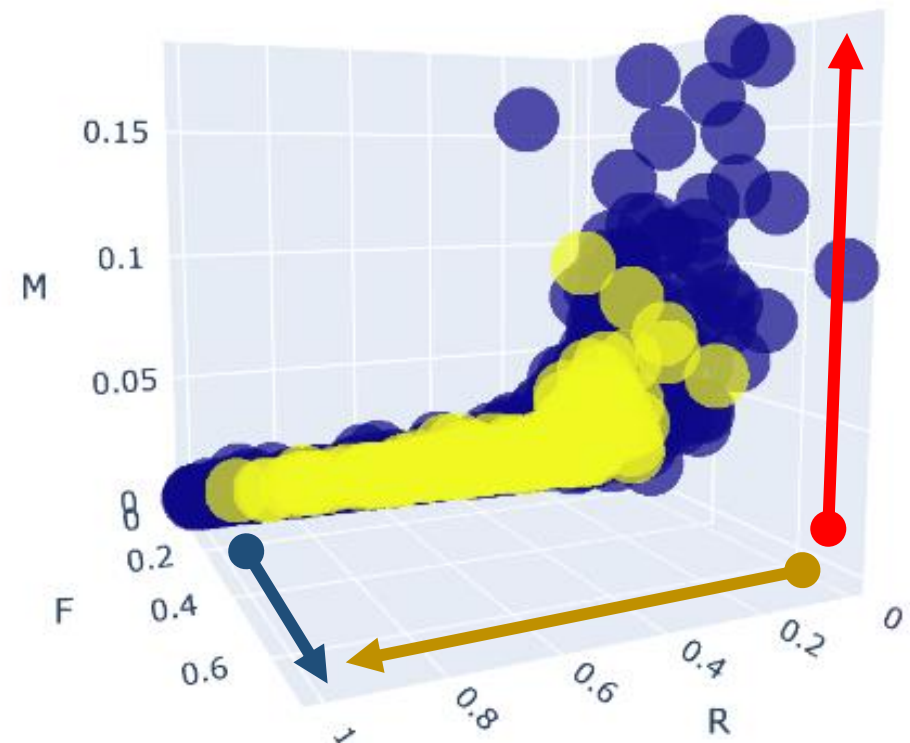
## 4.3 신규 활동고객 분석 (1) 신규 고객의 RFM 특성

우량 가망 고객군이 다수 존재하지만 **이탈 위험 고객이 대부분**  
⇒ **이탈을 막는 것이 우선**이며 그 후 우량 가망 신규 고객군을 잡을 전략이 필요함

R - F - M 별 고객 특성



R - F - M 에 대한 그래프



\*노란색 : 신규 고객군  
\*MinMax scaling한 결과를 plot하였음  
\*R의 경우 작을 수록 최근에 방문



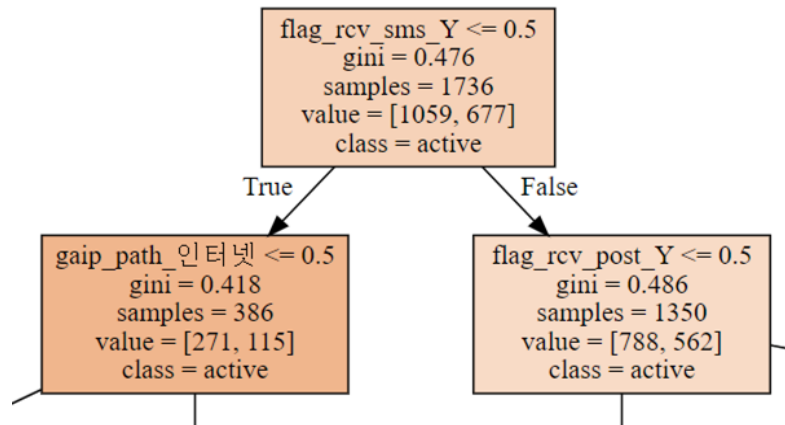
## 4.3 신규 활동고객 분석 (2) 분류 모델 - 의사결정나무

신규 고객 데이터를 이용하여 의사결정나무 분석을 실시  
활동고객의 핵심 인자로 추정되는 "**SMS 수신 여부**"를 발견

변수 중요도

변수	중요도
SMS 수신 동의	0.241
POST 수신 동의	0.180
50대	0.102
30대	0.100

분류 모델 시각화 (상위)



분류 결과

신규고객		예측	
		비활동	활동
실제	비활동	456	23
	활동	235	30

- 모델 정확도: 0.653
- 특이도: 0.95 (실제 비활동 고객을 비활동 고객으로 분류할 확률)

### 4.3 신규 활동고객 분석 (3) 분류 모델 - 랜덤포레스트

랜덤포레스트 분석 결과, "**SMS, email 수신 여부**"가 중요한 변수임을 확인할 수 있음.  
신규고객들의 최초 구매 유도하기 위해 SMS, email 과 관련된 프로모션을 진행할 예정

변수 중요도

변수	중요도
email 수신 동의	0.278
SMS 수신 동의	0.154
SMS 수신 거부	0.141
email 수신 거부	0.090

분류 결과

신규고객		예측	
		비활동	활동
실제	비활동	448	9
	활동	276	11

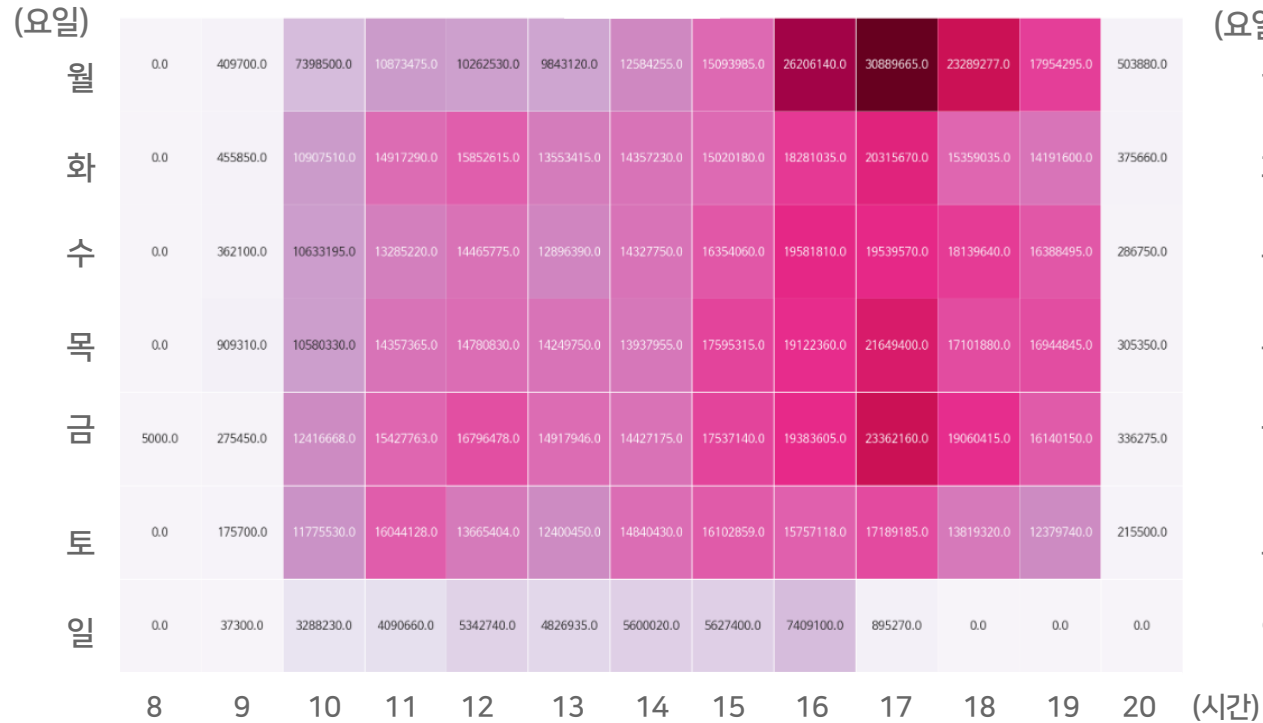
- 모델 정확도: 0.617
- 특이도: 0.98 (실제 비활동 고객을 비활동 고객으로 분류할 확률)

## 4.4 시간에 따른 매장별 매출 분석 목적

시간에 따른 매장별 매출 및 제품 분석

## 4.4 시간에 따른 매장별 매출 분석 (1) 매장 / 요일 / 시간대 매출 분석 - 오프라인

제주, 서귀포 매장에서 점심, 저녁 준비 시간에 매출액이 급증



제주 매장의 요일-시간대별 매출액

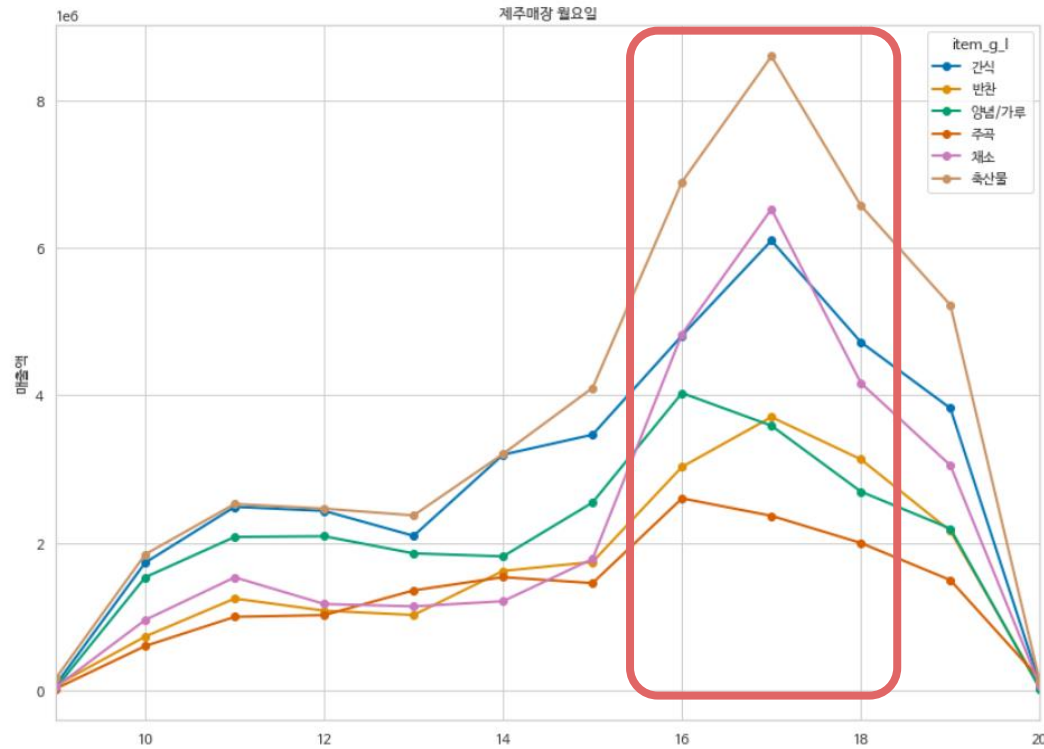


서귀포 매장의 요일-시간대별 매출액

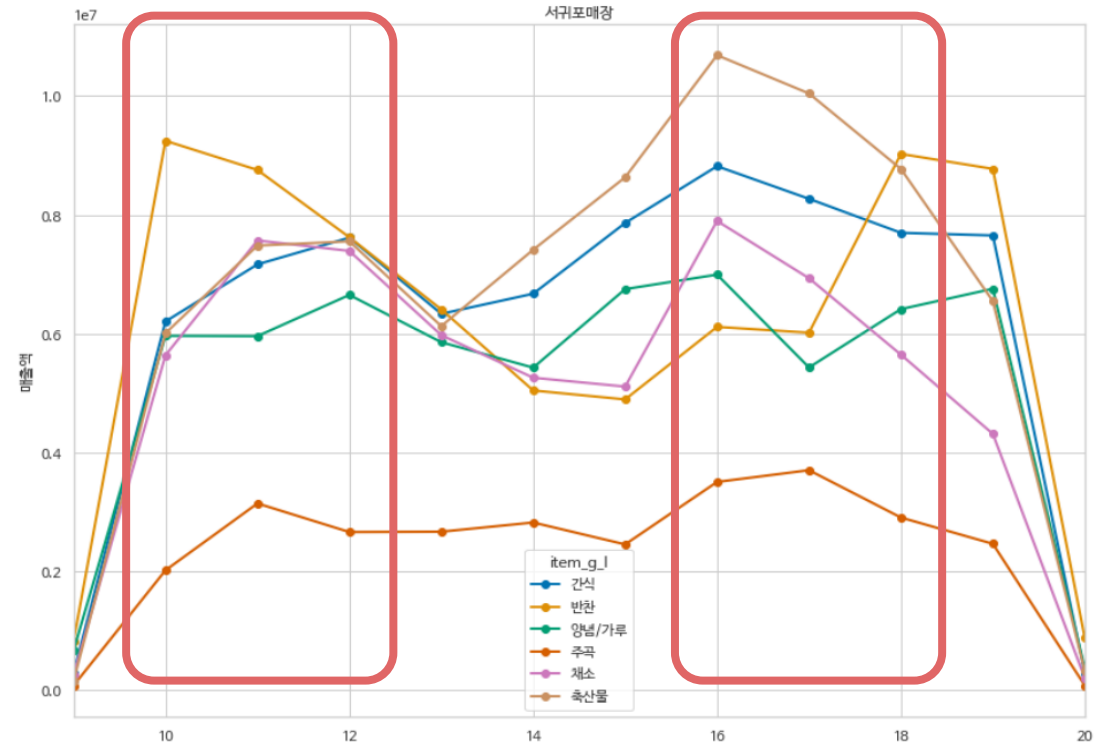
- 제주 매장은 15시부터 18시 사이, 서귀포 매장은 월요일에서 금요일 10시부터 12시, 16시부터 18시 사이에 매출액이 급증하기 때문에 매출액이 가장 큰 제품을 분석하여 재고 관리 및 인력 효율화등의 정책이 필요

## 4.4 시간에 따른 매장별 매출 분석 (2) 매출이 높은 요일의 시간대별 대분류 매출 분석 - 오프라인

점심, 저녁 시간대에 축산물과 반찬의 매출액이 큰 폭으로 상승



제주 매장의 시간대별 대분류 매출액 (월요일)

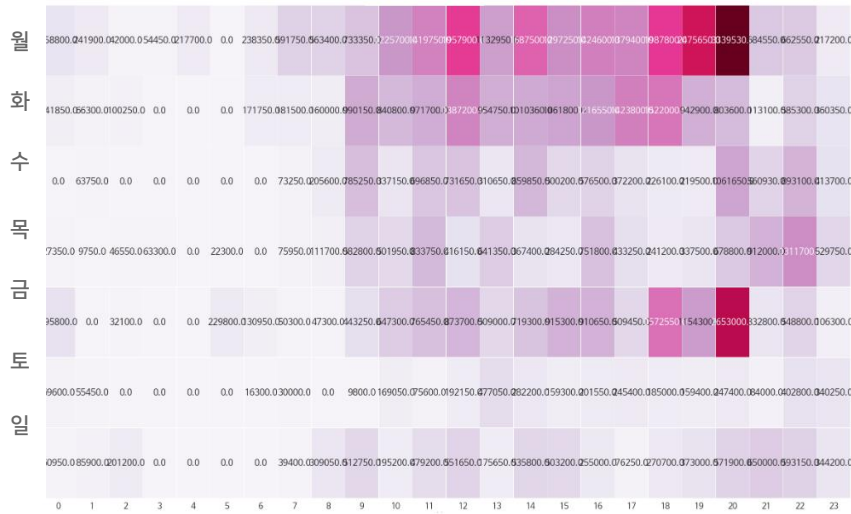


서귀포 매장의 시간대별 대분류 매출액(일주일)

- 제주 매장의 경우 월요일 16시~18시에 축산물, 서귀포 매장의 경우 10시~12시에 반찬, 15시~18시에 축산물 매출 급증
- 따라서 해당 시간 이전에 관련 제품들의 재고 관리등과 같은 운영 프로세스 개편을 통한 효율화 상승이 필요

## 4.4 시간에 따른 매장별 매출 분석 (3) 매장 / 요일 / 시간대 매출 분석 - 온라인

온라인 매장에서는 월요일, 목요일에 매출이 과도하게 집중



인터넷 주문에 의한 요일-시간대별 매출액



모바일 주문에 의한 요일-시간대별 매출액

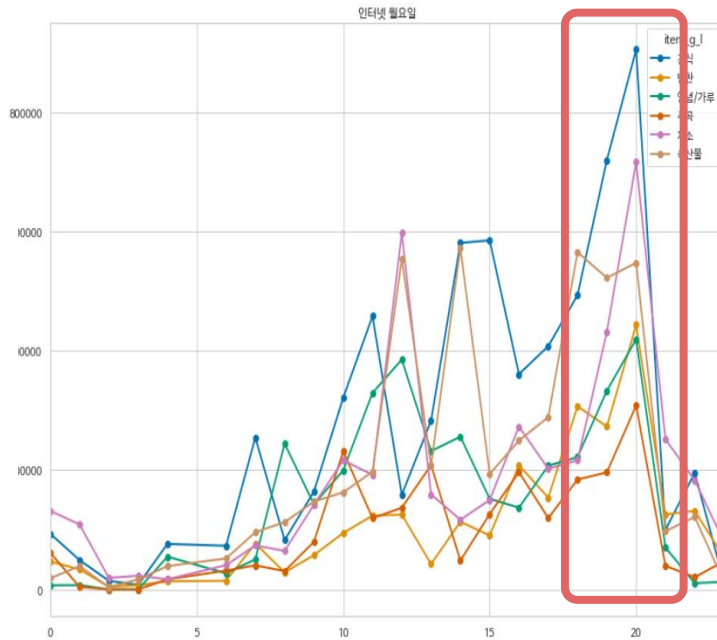


전화 주문에 의한 요일-시간대별 매출액

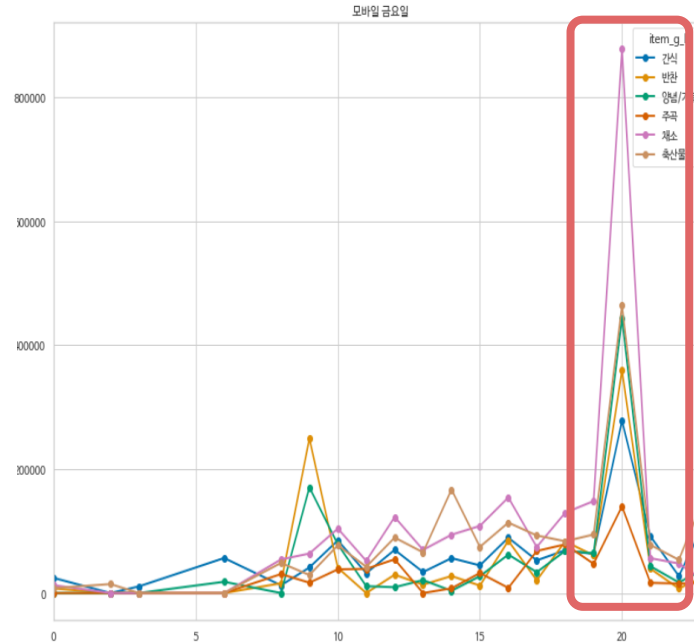
- 오프라인과 다르게 온라인 주문은 특정 요일 특정 시간에만 극히 몰려있음
- 인터넷 주문, 전화 주문의 경우 월요일과 화요일, 모바일 주문은 목요일 특정 시간대에 주문이 급증하는 품목을 조사하여 배송 준비, 재고 관리등에 대한 운영 프로세스 조정 필요

## 4.4 시간에 따른 매장별 매출 분석 (4) 매출이 높은 요일의 시간대별 대분류 매출 분석 - 온라인

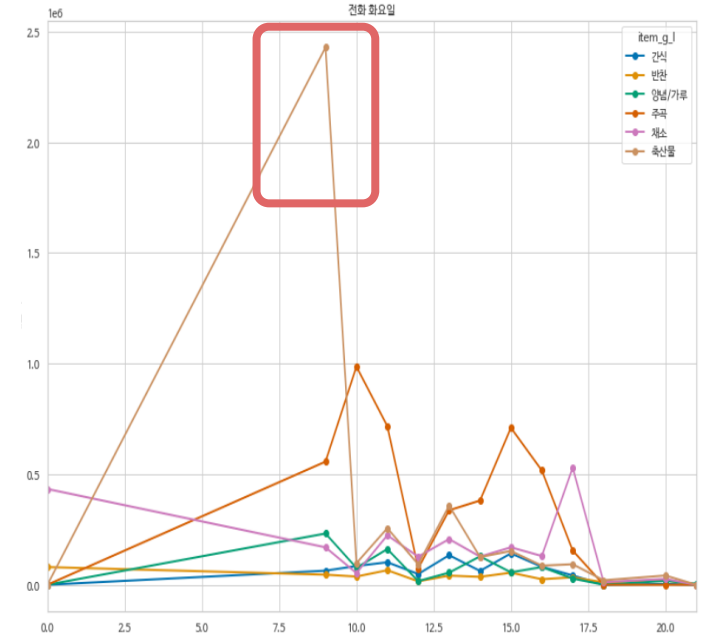
- 인터넷과 모바일 주문에서는 20시에 간식과 채소의 매출이 큰 폭으로 증가
- 전화 주문에서는 오전에 축산물의 매출이 큰 폭으로 증가



인터넷 주문-시간대별 대분류 판매 횟수 (월요일)



모바일 주문-시간대별 대분류 판매 횟수 (금요일)



전화 주문-시간대별 대분류 판매 횟수 (화요일)

- 인터넷 주문은 월요일 20시에 간식과 축산물, 모바일 금요일의 경우 20시에 채소의 매출이 급증
- 전화 주문은 화요일 오전에 축산물의 매출이 급증
- 수요 급증 시간에 관련된 품목들에 대한 배송 준비 및 재고 관리를 통한 운영 프로세스 개선 필요

## 개선안 도출

5.1 (1) 고객 세분화 및 고객 관리 방안 수립

(2) 등급별 할인쿠폰 및 포인트 적립

5.2 (1) 소풍 패키지 상품 기획

(2) 쿡으로 만든 제품 확보 및 소개

5.3 첫 구매 할인 쿠폰 발송 프로모션





5.4 (1) 오프라인 매장 재고 관리 및 판매 인원 배치

(2) 온라인 예측 배송 시스템 구축




## 5.1 개선안 적용방안 (1) 고객 세분화 및 고객 관리 방안 수립

등급 이동을 총 5가지로 분류하고 각각에 맞게 고객 세분화 전략 구상  
**ROYAL** 등급을 유지하는 고객들은(①) 평생 고객으로 만들기 위한 방안 수립  
 등급이 상승한 고객들은(②, ③) 충성고객으로 유도하기 위한 방안 수립

	① ROYAL 등급 유지	② 3등급 상승	③ 1, 2등급 상승
고객 특성	<ul style="list-style-type: none"> <li>· 충성도가 높은 고객</li> <li>· 고객 유지 및 서비스 비용이 낮고 수익성이 높음</li> </ul> 	<ul style="list-style-type: none"> <li>· 수익 향상 가능성이 높은 고객이므로 방문 유도가 필요</li> <li>· 아직 충성도가 높지 않으므로 세심한 고객 관리가 필요</li> </ul>	<ul style="list-style-type: none"> <li>· 최근 구매한 상품에 관심이 많은 고객</li> <li>· 만족도가 상승한 고객</li> </ul>
관리방안	<ul style="list-style-type: none"> <li>· <b>정기 배송 서비스</b> <ul style="list-style-type: none"> <li>- 반복적으로 구매하는 상품을 정기 배송으로 편리하게 받아볼 수 있는 서비스</li> <li>- 정기 배송 기간: 1년, 해당 제품 5% 할인</li> <li>- 지정한 날짜에 상품을 매달 배송함으로써 매번 같은 제품을 반복 구매하던 <u>고객의 불편함을 해소할 수 있음.</u></li> </ul> </li> <li>· <u>매장 입장에서는 지속적인 수익을 기대할 수 있음</u></li> </ul>	<ul style="list-style-type: none"> <li>· <b>기념일 선물 증정</b> <ul style="list-style-type: none"> <li>ex) 뷔페 식사권, 지역사회 공연 티켓 (클래식 음악회)</li> </ul> </li> <li>· <b>신상품 체험 기회 부여</b></li> </ul> 	<ul style="list-style-type: none"> <li>· 최근 구매한 상품에 대한 <b>만족도 조사 실시 (참여시 포인트 2배 적립)</b> </li> <li>· 만족도가 높은 제품 위주로 <b>상품 광고 문자 전송</b> </li> </ul>

## 5.1 개선안 적용방안 (1) 고객 세분화 및 고객 관리 방안 수립

등급이 하락한 고객을(④) 대상으로는 이탈을 방지하기 위한 전략 수립  
이탈한 고객을(⑤) 대상으로는 이탈 원인을 파악하여 행복그린 매장의 문제점 개선

	④ 2단계 하락	⑤ 이탈 고객 관리
고객 특성	<ul style="list-style-type: none"> <li>· 이탈 가능성이 있는 고객</li> <li>· 고객 만족도가 하락</li> </ul>	<ul style="list-style-type: none"> <li>· 매장 혹은 상품에 대한 불만으로 탈퇴한 고객</li> </ul>
관리방안	<ul style="list-style-type: none"> <li>· 고객 별 구매 상품 목록 및 주요 구매 시간 분석</li> <li>· 과거 구매한 상품에 대해 어떤 불만이 있었는지 설문 조사 실시(참여시 아메리카노 쿠폰 제공) </li> <li>· 구매 가능성이 높은 시간대 1시간 전, 할인 상품 홍보 문자 전송</li> </ul>	<p>이탈 고객의 이탈 원인 3가지</p> <ul style="list-style-type: none"> <li>① 실망이탈: 자사 제품 혹은 서비스 대한 실망</li> <li>② 경쟁이탈: 경쟁사의 공격적 마케팅(저렴한 가격)</li> <li>③ 자격이탈: 불가피한 사유(거주 이전, 경제적 이유)</li> </ul> <p>이탈한 고객이 3가지 원인 중 어떤 요인으로 탈퇴했는지 설문조사(참여시 아메리카노 쿠폰 제공)로 파악하여 행복그린 매장의 문제점 개선</p>

## 5.1 개선안 적용방안 (3) 고객 특성별 관리 방안










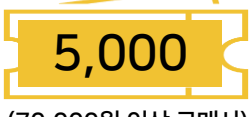



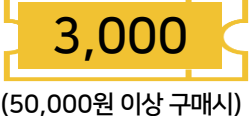






군집 특성에 맞는 적절한 "유지 전략" 을 도입한다.

군집	R 점수	F 점수	M 점수	군집 특성
자주 방문하지도 않으며 최근에 방문하지도 않은 고객	3	1	2	이탈 위험 고객
최근에 방문했지만 자주 오지 않아 소비 금액이 적은 고객	4	1	2	관심 고객
일반 고객	4	2	3	일반 고객
최근에 방문했으며 많이 소비했지만 자주 오지는 않는 고객	5	2	4	충성 고객으로 전환 가능한 고객
R,F,M이 모두 높은 충성 고객	6	3	5	충성 고객

- 이탈 위험 고객 : 휴면 계정 복구 이벤트와 같은 행사를 통해 다시 방문할 수 있도록 유도
- 관심 고객 : 한 번 방문마다 10,000원 이상 구매 시 스탬프를 찍어주고 총 10번 방문 시 무료 배송 쿠폰 제공  
금액을 늘리는 것보다는 자주 방문할 수 있도록 유도하여 매장에 대한 관심도를 높여야 함.
- 충성 고객으로 전환 가능한 고객 : 자주 방문할 수 있도록 관심 상품에 대한 할인 정보를 문자로 제공
- 충성 고객 : 지속적인 특별 혜택을 통해 계속해서 매장의 충성 고객이도록 유지하는 전략이 필요

## 5.1 개선안 적용방안 (2) 등급별 할인 쿠폰 및 포인트 적립

포인트 적립, 무료 쿠폰 등의 멤버십 서비스는 고객의 재구매를 유도하고  
단골 고객을 확보하여 매출 확대에 기여하는 효과가 있을 것이라 기대

고객 등급	 ROYAL	 PLATINUM	 GOLD	 SILVER
등급별 할인 쿠폰	 10,000 (100,000원 이상 구매시)	 7,000 (100,000원 이상 구매시)	 6,000 (100,000원 이상 구매시)	 5,000 (100,000원 이상 구매시)
	 7,000 (70,000원 이상 구매시)	 5,000 (70,000원 이상 구매시)	 4,000 (70,000원 이상 구매시)	 3,000 (70,000원 이상 구매시)
	 5,000 (50,000원 이상 구매시)	 3,000 (50,000원 이상 구매시)	 2,000 (50,000원 이상 구매시)	 2,000 (50,000원 이상 구매시)
포인트 적립	 5 %	 3 %	 1 %	 0.5 %

## 5.2 개선안 적용 방안 (1) 소풍 패키지 상품

날씨가 좋은 날에 '스낵' 제품을 구매하는 고객들에게 '김가공' 제품도 함께 판매하여 매출 증대를 기대

날씨에 따른 패키지 상품 판매



SMS, email을 통한 제품 추천

**7월 행복이벤트**

우리 농산물로 건강도 챙기고 농업인에게 힘도 실어주세요

해남미소 8주년 기념 할인행사 & 구매고객 선물 폭탄

농수특산물 최고 30% 할인

양품 김치세트, 군고구마달랑이, 양품 청채소, 쌀, 잡곡, 소시지 등

**30%**

농산물 구매하고, 선물은 받으세요!

- 대상: 7월 1일 ~ 10일까지 구매한 모든 고객 (주최를 통해 선물이 발송)
- 당첨선물
  - MVP: 8명 / 당첨품 10kg
  - VIP: 8명 / 당첨품 4kg
  - 당첨품: 8명 / 해남미소 기념품 (구매상품 후기 및 응원메시지 작성)
- 주최일: 7월 26일
- 당첨자 확인: 해남미소 홈페이지

- 미세먼지 농도가 낮고 따뜻한 날씨에 '스낵'과 '김가공' 등 소풍과 관련된 상품들을 패키지로 묶어서 판매
- 휴일 또는 날씨가 좋은 날에 SMS, email을 통해 야외 활동을 유도하는 이미지 및 문구와 함께 '스낵'과 '김가공' 제품에 대한 상품 추천과 할인 정보를 소개

## 5.2 개선안 적용 방안 (1) 소풍 패키지 상품

**"손수 만드는 소풍 도시락" 패키지 : 김밥 재료 + 돈가스 재료 + 도시락 반찬으로 구성**

소풍 도시락을 위해 필요한 재료를 패키지로 판매하거나 함께 진열

1. 돈가스와 관련된 상품
  - 돈가스 소스, 돈가스용 고기, 빵가루, 현미유
2. 김밥과 관련된 상품
  - 긴 어묵, 네모난 햄, 단무지, 전장김, 김밥용 김
3. 도시락 반찬과 관련된 상품
  - 소시지 볶음, 어묵 볶음 등 도시락에 자주 사용되는 반찬
4. 작은 용량의 소스류 구비
  - 일회성으로 사용할 확률이 큰 돈가스 소스나 김밥용 단무지 등은 작은 용량을 준비하여 구매에 부담이 없도록 조절
5. 함께 곁들일 계절 과일을 함께 추천



## 5.2 개선안 적용 방안 (2) 콩으로 만든 제품을 선호하는 고객층을 위한 프로모션

'두부' 제품을 구매하는 고객들에게 여러 가지 '곡식가루' 제품 또는 콩으로 만든 제품을 소개하고 구매를 유도하면서 매출 증대를 기대

콩으로 만든 제품 확보



'두부' 제품 근처에 시식 코너 설치



- 콩으로 만든 다양한 제품들을 매장 내에 확보하고 email, 우편을 통해 콩에 대한 효능을 소개하면서 다양한 '두부', '곡식가루' 제품을 추천
- '두부' 제품 근처에 콩국, 콩국수와 같은 제품을 시식해볼 수 있는 코너를 설치하여 '곡식가루' 또는 콩으로 만든 제품의 구매를 유도



## 5.2 개선안 적용 방안 (2) 된장찌개/청국장 패키지 상품

**“오늘같은 날은 된장찌개 ! ” 패키지 : 순두부 + 들깨가루/생콩가루 + 된장/청국장**

순두부가 주로 팔리는 오전 10시, 오후 6시에 그 날의 식사 메뉴를 추천

1. 순두부만 구매하는 고객에게  
“더 고소한 맛을 위한 들깨가루는 어떠세요?”  
“더 깊은 맛을 위한 생콩가루는 어떠세요?”
2. 버섯, 고추 등 관련 식재료를 함께 묶어  
패키지로 판매





## 5.3 개선안 적용 방안 - 첫 구매 할인쿠폰 발송 프로모션

새롭게 유입된 신규고객들이 SMS, email 수신 동의 시,  
SMS로 첫 구매 할인쿠폰을 발송하여 홍보채널 확보 및 최초 구매를 유도

### 첫 구매 할인 마케팅 예시

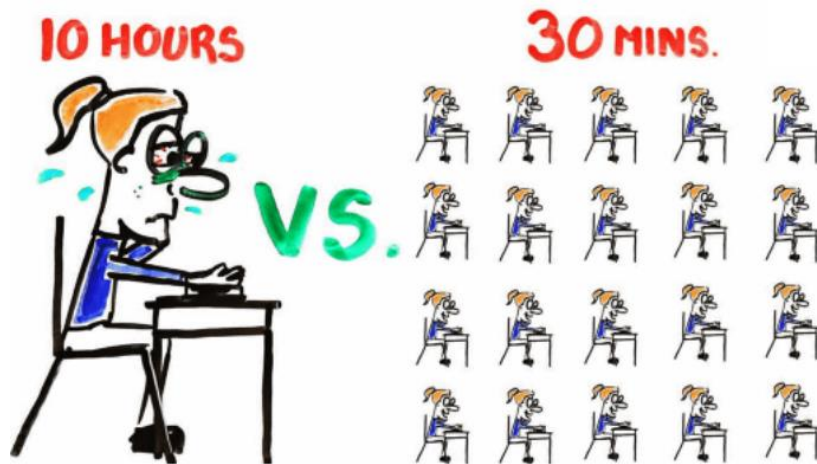


- SMS와 email은 추후에 제품 할인 이벤트 공지나 고객관리 차원의 채널로도 활용이 가능

## 5.4 개선안 적용 방안 - 효율적인 인원 배치 및 예측 배송 시스템

시간별 수요가 많은 제품에 대한 효율적인 인원배치 및 재고 관리와  
예측 배송 시스템을 통해 원가 절감을 기대

[ 오프라인 매장 ] 재고 관리 및 판매 인원 배치



[ 온라인 매장 ] 예측 배송 시스템



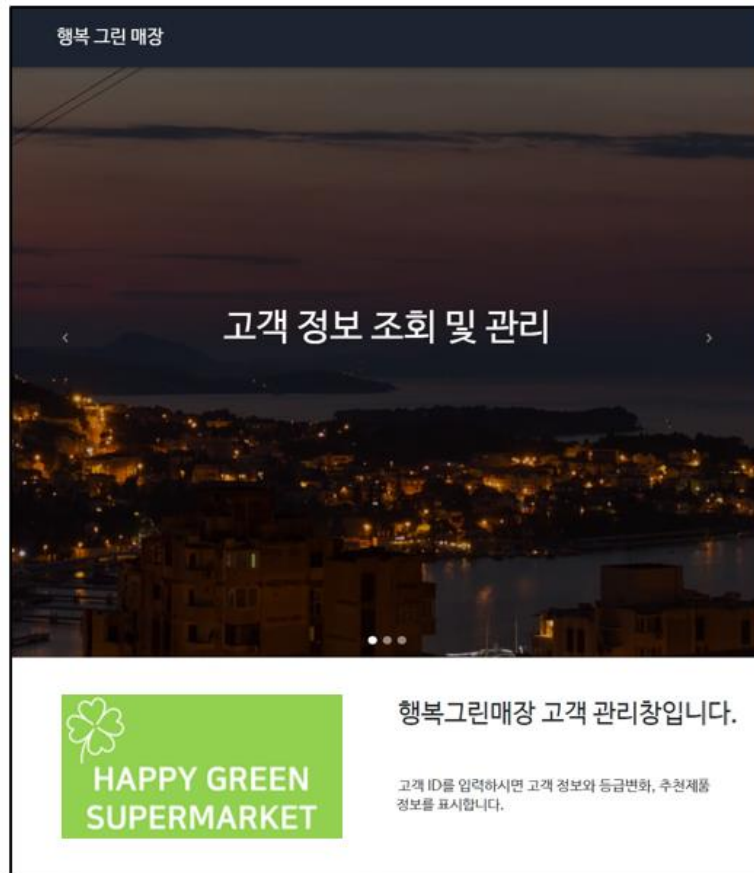
- (오프라인 매장) 구매 파워를 향상시키고 재고 관리를 통해 원가절감을 도모  
ex) 반찬이 많이 팔리는 오전 시간대에 반찬 시식 등 판촉을 위한 직원 추가 배치
- (온라인 매장) 매장의 입장에서는 효율적인 재고 및 인력 관리, 고객의 입장에서는 빠른 배송 서비스 체험 가능

## 결론

1. 등급 이동에 따른 혜택과 프로모션을 통해 등급별 평균 객단가가 10% 상승할 것으로 기대  
⇒ **68,364,000 원**
2. 연관 분석을 통하여 김가공과 곡식가루의 판매량을 2배 이상 늘릴 수 있을 것으로 기대  
⇒ **5,193,600 원**
3. 신규고객 최초 구매 유도를 통해 신규활동고객 전환율을 50% 이상 증가할 수 있을 것으로 기대  
⇒ **84,150,000 원**
4. 매장 운영 방안 개선과 배송 시스템의 개선을 통해 장기적인 경쟁력을 갖출 수 있을 것으로 기대  
⇒ **19,200,000 원**

**" 총 1.7억원의 매출 증대를 기대할 수 있음 "**

# 행복 그린 매장 고객 관리 사이트



메인 화면

고객 분석

고객 ID

고객 ID를 입력하세요.

검색

결과

고객 ID: C472474

나이: 51세

가입일: 2016-12-28

주소: 서귀포시 동홍동

현재 등급: Royal

2018년 대비 2019년 등급 변화: 무

<등급별 관련 프로모션>

정기 배송 서비스 제공

<가장 최근 구매 내역>

2019-12-30: [시금치, '봄동', '돼지가공', '한우', '떡', '두부'] 구매

<서귀포매장 50대 연관 제품 추천>

[돼지가공 → 김가공, '두부' → 곡식가루]

고객 정보 조회 및 연관제품 추천

## 6. Learned Lesson

분석 초반에는 RFM을 이용한 고객 특성 파악 및 등급 분류를 담당하여 분석했다. 사용 금액, 방문 빈도, 최근성을 도입하여 세가지 측면에서 고객을 평가해보았는데, 어떻게 점수를 산출하고 등급을 나눌 것인지에 대해 관련 논문을 찾아보며 직접 구현해보고 비교해보는 과정이 재미있었다. 또한 여러가지 방법 중 어떤 것을 사용할지에 대해 팀원들과 여러 번 토의를 하였는데, 각자 생각하지 못한 부분을 서로 채워가며 진행할 수 있어 좋았다.

고객 분석이 끝난 후에는 요일-시간별 선호 제품을 분석해보았다. 많은 통계량과 그래프 사이에서 유의미한 결과를 찾아내는 것이 어려웠다. 팀원들에게 코드를 공유하며 함께 고민해보니 제각기 다른 시각으로 해석해왔고 토론하며 의견을 나누는 시간에 가장 많이 성장한 것 같다.

다양한 방법으로 데이터를 들여다보고자 Rodgers모형 대신 K-means를 이용해보거나 연관 규칙 대신 item2vec을 이용해보았다. 비슷한 결과가 나오기도, 새로운 결과가 나오기도 했는데 여러 방향으로 바라보며 해석하고 비교해보는 과정이 굉장히 흥미로웠다.

수많은 통계량과 그래프, 파생 변수 사이에서 유의미한 결과를 찾아낼 수 있는 능력이 아주 중요함을 깨달았다.  
탐색적 분석을 통해 데이터를 충분히 이해해야 한다.  
데이터에 대한 통찰력과 전달력을 기를 수 있도록 노력할 것이다.

POSCO AI / BIGDATA ACADEMY 9TH 빅데이터 프로젝트

# THANK YOU

[ A반 1조 ] 강민구 권수민 이용현 이정우 안태윤 정해유

## 7. 첨부 - RFM 등급을 통한 등급 상/하향화 예시

### [ 등급 상향 조정 예시 ]

자주 방문하며 큰 금액을 쓰지만 하위 등급으로 분류되었던 고객들이 상향 조정 되었음

고객 ID	기존 등급	R	F	M	새로운 등급
C418404	Gold	2	70	1,581,000	Royal
C404906	Gold	4	79	1,821,460	Royal
C422720	Gold	5	113	4,075,630	Royal
C541996	Gold	3	84	1,866,885	Royal
C472947	Gold	2	70	1,914,550	Royal

- 1년동안 113번(약 3일에 1번꼴로 방문)하여 총 400만원을 지출한 고객이 기존에는 하위 등급인 Gold로 분류되어 있었음.

## 7. 첨부 - RFM 등급을 통한 등급 상/하향화 예시

### [ 등급 하향 조정 예시 ]

자주 방문하지 않으며 총 사용 금액 역시 낮은 고객들이 하향 조정 되었음

고객 ID	기존 등급	R	F	M	새로운 등급
C222499	Royal	75	1	20,350	Silver
C354460	Royal	145	1	21,200	Silver
C450409	Royal	107	4	67,100	Silver
C450436	Royal	187	1	15,450	Silver
C364526	Royal	80	4	69,100	Silver

- 1년 동안 한 번만 방문하여 1만원을 썼음에도 Royal로 분류되어있던 고객을 Silver 등급으로 하향 조정하게 되었음.



## 7. 첨부 - 제주 매장 30대에 대한 연관규칙 분석

### [ 제주 매장 30대 고객층 연관규칙 ] '스낵' → '김가공'

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
78	(스낵)	(빵)	0.294330	0.129124	0.042912	0.145797	1.129125	0.004907	1.019519
15	(스낵)	(김가공)	0.294330	0.133119	0.050129	0.170315	1.279425	0.010948	1.044832
35	(스낵)	(돼지가공)	0.294330	0.110180	0.039820	0.135289	1.227886	0.007390	1.029037
53	(스낵)	(두부)	0.294330	0.190851	0.055670	0.189142	0.991047	-0.000503	0.997893
95	(스낵)	(우유/산양유)	0.294330	0.108247	0.035696	0.121278	1.120382	0.003835	1.014830

- 제주매장 30대 고객들이 '스낵'을 구매할 때 위와 같은 제품들을 함께 구매하는 경향을 보이는데 모든 규칙에 대해 프로모션을 진행할 수 없으므로 가장 유용한 규칙을 선택하여 프로모션을 진행하기로 결정
- 신뢰도(confidence)는 항목집합 X(antecedents)를 포함하는 거래 중에서, 항목집합 Y(consequents)도 포함하는 거래 비율이므로 신뢰도가 높을수록 유용한 규칙일 가능성이 높음
- '스낵'→'두부' 연관규칙은 향상도(lift)가 1보다 작은 음의 상관관계를 보이므로 유용한 규칙이라고 판단하지 않음
- 신뢰도도 높고 향상도도 1보다 큰 '스낵'→'김가공' 연관규칙에 대해 살펴봐야함

## 7. 첨부 - 서귀포 매장 50대에 대한 연관규칙 분석

### [ 서귀포 매장 50대 고객층 연관규칙 ] '두부' → '곡식가루'

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(두부)	(곡식가루)	0.317570	0.120174	0.074187	0.233607	1.943910	0.036023	1.148009
23	(두부)	(곡식가루, 장류/기타)	0.317570	0.053145	0.040781	0.128415	2.416304	0.023904	1.086360
7	(두부)	(된장)	0.317570	0.063557	0.040781	0.128415	2.020459	0.020597	1.074414
8	(두부)	(면/기타)	0.317570	0.102169	0.071800	0.226093	2.212926	0.039355	1.160127
10	(두부)	(장류/기타)	0.317570	0.097397	0.056399	0.177596	1.823421	0.025469	1.097517
17	(두부)	(곡식가루, 면/기타)	0.317570	0.067896	0.054230	0.170765	2.515102	0.032668	1.124053

- 서귀포매장 50대 고객들이 '두부'를 구매할 때 위와 같은 제품들을 함께 구매하는 경향을 보이는데 모든 규칙에 대해 프로모션을 진행할 수 없으므로 가장 유용한 규칙을 선택하여 프로모션을 진행하기로 결정
- 신뢰도(confidence)는 항목집합 X(antecedents)를 포함하는 거래 중에서 항목집합 Y(consequents)도 포함하는 거래 비율이므로 신뢰도가 높을수록 유용한 규칙일 가능성이 높음
- 신뢰도가 가장 높고 향상도도 1보다 큰 '두부'→'곡식가루' 연관규칙에 대해 살펴봐야함

## 7. 첨부 - 예상 매출액 계산 방법

1. 등급 이동에 따른 혜택과 프로모션을 통해 등급별 평균 객단가가 10% 상승할 것으로 기대

$$\Rightarrow 683,642,911 \text{ (2019년 총 매출액)} \times 0.1 \text{ (기대 상승율 10\%)} = \underline{68,364,291 \text{ 원}}$$

2. 연관 분석을 통하여 김가공과 곡식가루의 판매량을 2배 이상 늘릴 수 있을 것으로 기대

$$\begin{aligned} \Rightarrow & 690 \text{ ('김가공' 제품 증가수)} \times 3280 \text{ ('김가공' 제품 평균가격)} + 330 \text{ ('곡식가루' 제품 증가수)} \times 8880 \text{ ('곡식가루' 제품 평균가격)} \\ & = 2,263,200 + 2,930,400 = \underline{5,193,600 \text{ 원}} \end{aligned}$$

3. 신규고객 최초 구매 유도를 통해 신규활동고객 전환율을 50% 이상 증가할 수 있을 것으로 기대

$$\Rightarrow 550 \text{명 (2020년 예상 신규 활동고객 인원)} \times 153,000 \text{ 원 (19년 신규 고객 객단가)} = \underline{84,150,000 \text{ 원}}$$

4. 매장 운영 방안 개선과 배송 시스템의 개선을 통해 장기적인 경쟁력을 갖출 수 있을 것으로 기대

$$\Rightarrow 80 \text{만원 (금요일 20시 증가된 채소 매출액)} * (4 \text{주}) * (12 \text{달}) = \underline{19,200,000 \text{ 원}}$$